

LERNPROZESSE IN NICHT – MARKOFF'SCHEN
STOCHASTISCHEN AUTOMATEN

Rainer B. Pelka

1975



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41551910_0020

Dissertationsdruck

Blasaditsch GmbH • 89 Augsburg 17 • Theodor-Heuß-Platz 3 • Telefon (0821) 51 85 63

LERNPROZESSE IN
NICHT - MARKOFF'SCHEN
STOCHASTISCHEN AUTOMATEN

Inaugural - Dissertation
zur Erlangung des Doktorgrades
der Ludwig-Maximilians-Universität
zu München

Fachbereich: Psychologie



vorgelegt von
Rainer B. Pelka in München

Dissertationsdruck W. Blasaditsch, 89 Augsburg

1975



Referent : Prof. Dr. W. Butollo
Koreferent: Prof. Dr. H. Schiefele

Tag der mündlichen Prüfung: 31. Juli 1975

INHALTSVERZEICHNIS

EINFÜHRUNG	I
------------------	---

Kap. I: ALLGEMEINES MODELL EINES LERNSYSTEMS

1.1 Grundannahmen für ein Verhaltenssystem.....	1
1.2 Homöostatische Automaten und Lernsysteme.....	10
1.3 Allgemeine Eigenschaften des stochastischen Lernprozesses.....	19

Kap. II: GRENZWERTSÄTZE ÜBER LERNPROZESSE

2.1 Konvergenzkriterien für unabhängige Lernsysteme.....	26
2.2 Approximationsbedingungen für abhängige Lernsysteme.....	39
2.3 Strukturapproximation in differentiellen Lernautomaten.....	43
2.4 Modelle mehrstufiger Lernsysteme.....	50

*

Kap. III: ERLÄUTERUNG DES FORMALEN MODELLS

3.1 Analyse der Lernphänomene.....	53
3.1.1 Das Lernsystem als Automat.....	55
3.1.2 Der Lernprozeß.....	58
3.1.3 Das homöostatische Prinzip.....	59
3.1.4 Das Reinforcementprinzip.....	62
3.1.5 Das Differenzierungs- und Ökonomieprinzip.....	63
3.1.6 Transfereigenschaften.....	67
3.1.7 System der Sollwerte.....	68
3.2 Diskussion der Grundannahmen für ein Lernsystem...	73
3.3 Konsistenz des Modells mit empirisch gestützten Lernphänomenen.....	76

Kap. IV: ANWENDUNGSMÖGLICHKEITEN IN DER EMPIRISCHEN
PSYCHOLOGIE

4.1	Koppelung von Systemen	79
4.2	Vorhersageprobleme im pädagogischen Bereich	82
4.2.1	Auslese	84
4.2.2	Ausbildung	85
4.2.3	Plazierung	86
4.2.4	Prognose	87
4.2.5	Prognostische Diagnose	88
4.3	Anwendungen in der Verhaltenstherapie	89
4.4	Computersimulationen von Lernprozessen und Vergleich mit empirischen Befunden	90
4.4.1	Programmiertes Modell eines unabhängigen, differentiellen Lernsystems	90
4.4.2	Die Zahl der frei wählbaren Parameter	92
4.4.3	Flußdiagramm des Computerprogramms	93
4.4.4	Simulationsergebnisse im Vergleich mit psychologischen Experimenten	94
	ZUSAMMENFASSUNG der wichtigsten Ergebnisse	110
	LITERATURVERZEICHNIS	113

EINFÜHRUNG

Die bisher bekannten Versuche, im Rahmen der Psychologie mit formalisierten Theorien Lernprozesse bei Tieren oder Menschen zu erklären, verlangten erhebliche Kompromisse, die zwei der bekanntesten Lerntheoretiker, E.R.HILGARD und G.H.BOWER (1971) zu der Aussage veranlaßten: " Die Konstruktion einer völlig zufriedenstellenden Lerntheorie wird wahrscheinlich noch lange Zeit eine unvollendete Aufgabe bleiben."

Die von Psychologen konzipierten Lerntheorien konzentrieren sich in der Regel schwerpunktmäßig auf bestimmte Prozesse, während andere Lernphänomene - wenn überhaupt - nur flüchtig in den Rahmen des Modells eingebettet werden. Nahezu allen psychologischen Lerntheorien gemeinsam ist jedoch der geringe Formalismus bei dem Aufstellen der Postulate und ihrer Ableitungen, die eine Kategorisierung oder gar Quantifizierung der Beziehungen nahezu unmöglich machen. Dies erschwert zwar die Falsifizierbarkeit einzelner Theorien, aber auch ihre Anwendbarkeit.

In neuerer Zeit wurden zunehmend mathematisch orientierte Modelle von Lerntheorien aufgestellt, die die wünschenswerte Präzisierung besitzen, um empirische Überprüfungen zuzulassen. (vgl. BUSH / MOSTELLER ; 1955, ESTES; 1964, IOSIFESCU / THEODORESCU; 1969, FEICHTINGER; 1970, MENZEL; 1970, MENDEL / FU; 1971, NORMAN; 1972). Bei diesen Modellen wird jedoch eine erhebliche Reduktion der untersuchten Lernphänomene in Kauf genommen, um die besonders erforschten Bereiche der Wahrscheinlichkeitstheorie und mathematischen Statistik, insbesondere die Kenntnisse über Markoff'sche oder stationäre Prozesse verwenden zu können. Dies ermöglicht einerseits die Ableitung mathematisch anspruchsvoller Grenzwertsätze über das Verhalten der den Lernprozeß beschreibenden zufälligen Variablen, verlangt aber den Verzicht der Berücksichtigung von Lernphänomenen, die uns wesentlich für eine umfassende, psychologisch relevante Lerntheorie erscheinen.

Die vorliegende Arbeit kann nicht den Anspruch erheben, dieses Dilemma völlig zu beseitigen, sie versucht jedoch einen ersten Lösungsansatz aufzuzeigen. Die Tatsache, daß sich bereits in diesem Stadium für die psychologische Praxis interessante Ergebnisse ableiten lassen, ermutigt zu dem Versuch.

Unter Berücksichtigung der bereits im Rahmen psychologischer Untersuchungen bekannten Lernphänomene und unter Zuhilfenahme jener Theorien versuchten wir eine äußerst sparsame Modellbildung, die den experimentell gefundenen Tatsachen mathematisch formulierte Gesetzmäßigkeiten als Postulate oder Theoreme unterlegt. Das verlangte zunächst eine Bescheidung in der Verwendung mathematischer Mittel. Spezielle Untersuchungen - etwa im Rahmen der Automatentheorie oder im Bereich Markoff'scher Prozesse - konnten nicht herangezogen werden, weil ihre Voraussetzungen, wie in Kapitel III erläutert wird, nicht den inhaltlichen Anforderungen entsprachen, die wir an eine global anwendbare Lerntheorie zu stellen versuchten. Aus diesem Grund war es uns auch kaum möglich, Zeitschriftenliteratur neueren Datums in die Arbeit einzubeziehen. Soweit im einschlägigen Bereich mathematisch orientierte Publikationen gefunden wurden, gingen die Themen von so speziellen Voraussetzungen aus, daß sie für den von uns gewählten Ansatz nicht unmittelbar herangezogen werden konnten.

In dieser Arbeit versuchen wir, ein formalisiertes Modell darzustellen, das den Lernprozeß als einen stochastischen Prozeß und das Lernsystem als einen Automaten versteht. Dieser Prozeß ist nicht markoff'sch, stationär oder gar unabhängig, weil das (zu erlernende) Verhalten des Automaten von der Umwelt-Automat-Beziehung abhängig ist, die sich in nichtvorhersagbarer Weise verändern kann. Somit können wir weder die Übergangswahrscheinlichkeiten noch die bedingten Verhaltenswahrscheinlichkeiten als konstant annehmen. Natürlich müssen bestimmte Annahmen über die bedingten Reaktionswahrscheinlichkeiten der Umwelt gemacht werden, damit "vernünftige" Lernprozesse erwartet werden können, aber die Sätze in Kapitel II zeigen, daß jene schwächer sein können als dies bei Markoff'schen Prozessen der Fall ist.

Die unterschiedliche Ausdrucksweise in den Bereichen der Mathematik und der Psychologie zwang uns dazu, die Arbeit zweiteilig aufzubauen. Im ersten Teil (= Kapitel I und II) wird zunächst das Modell formal spezifiziert und nach den Definitionen des Homöostatischen Systems (= HS) und des Lernsystems (=LS) werden die mathematisch wesentlichen Grenzwertsätze für unabhängige, abhängige sowie differentielle Lernsysteme abgeleitet. Mathematisch fundamental stützten wir uns dabei auf die Theorie der Semimartingale (vgl. DOOB, 1965).

Im zweiten Teil (= Kapitel III und IV) zeigen wir im Rahmen einer mehr psychologisch orientierten Sprechweise, daß das vorliegende Modell mit den psychologischen Vorstellungen über die Natur des Lernvorgangs in Individuen verträglich ist. Im Zusammenhang mit den in Kapitel II gefundenen Ergebnissen folgt daraus, daß das hier konzipierte Modell eine brauchbare Grundlage für eine weiterreichende bzw. genauere Theorie des Lernens liefern kann. Zugleich wird (allerdings heuristisch) festgestellt, an welchen Stellen das Modell in seinen Implikationen über die uns bekannten mathematischen Lerntheorien hinausgeht.

Das Lernsystem wird als ein stochastischer Automat eingeführt. Charakteristisch für ein lernendes Subjekt ist die Reaktion auf die Umwelt, die über Eingabesignale $\hat{X}_t$ (t = diskreter Zeitparameter) für das Individuum bedeutsam wird, über Ausgabesignale $\hat{Y}_t$. Dabei verändert sich der "inner" Zustand $\hat{Z}_t$ des LS. Zur Kenntnis der Gesetzmäßigkeiten, denen ein LS unterliegt, ist die Angabe der Abbildungen $f_t: \prod_{\tau < t} \hat{E}_\tau \rightarrow (\hat{X}_t, \hat{Z}_t)$ und $g_t: \prod_{\tau < t} \hat{E}_\tau * (\hat{X}_t, \hat{Z}_t) \rightarrow \hat{Y}_t$ erforderlich, wobei $\hat{E}_\tau = (\hat{X}_\tau, \hat{Z}_\tau, \hat{Y}_\tau)$. f_t heißt t -Überföhrungsfunktion und gibt die Veränderung der "Situation" (= Stimulus-Zustands-Paar) als Folge der Vorgeschichte, d.h. der "Ereignisse" $E_{t-1}, E_{t-2}, \dots$ der Vergangenheit wieder. g_t heißt Ergebnisfunktion und gibt den situationsbedingten Output des LS wieder. Notwendig für das "Lernen" eines Automaten ist die Bedingung, daß die verhaltenssteuernde Funktion g_t eine nichtkonstante Funktion der Vorgeschichte ist.

Das vorliegende Modell ist jedoch so konzipiert, daß von der Vorgeschichte möglichst wenig berücksichtigt werden muß. Ziel einer Verhaltensänderung ist die Auswahl (und ggfls. das Auffinden) einer Verhaltensalternative mit hoher Wahrscheinlichkeit, die bestimmten vorgegebenen Zielkriterien genügt. Als unsere "Lerntheorie" charakterisierend haben wir dabei folgende Prinzipien in das Modell einbezogen:

- (1) **H o m ö o s t a s i e** ist zielbestimmend, d.h. nicht ein bestimmter Stimulus, sondern der Zustand des Systems ist Grundlage einer Reaktion des LS. Es gibt genau einen Gleichgewichtspunkt unter allen Zuständen. Jede Abweichung des Ist-Zustandes davon wird als Ungleichgewicht registriert. Das System versucht, durch geeignete Verhaltensweisen wieder das Gleichgewicht zu approximieren. Der Stimulus beeinflusst lediglich die Auswahl aus den verfügbaren Alternativen.
- (2) **R e i n f o r c e m e n t** bezüglich der Ungleichgewichtsverringerung ist die Basis der Verhaltensmodifikation. Jede Situation s (= Stimulus-Zustands-Paar (x, z)) hat zwei Aspekte, einen verhaltenssteuernden (für das nachfolgende Verhalten) und einen verhaltensbewertenden (für das vorausgegangene Verhalten). Die Ungleichgewichtsänderung $\gamma(\hat{z}_t, \hat{z}_{t-1})$ des bestehenden gegenüber dem vorausgegangenen Zustand wird zur Grundlage der Verhaltensmodifikation, d.h. des Lernprozesses. Ungleichgewichtsverringerung (Belohnung, Lust) erhöht die bedingte Wahrscheinlichkeit des vorausgegangenen Verhaltens auf die vorausgegangene Situation, Ungleichgewichtsvergrößerung (Bestrafung, Unlust) verringert sie.
- (3) Wahrnehmungsraster und Verhaltensraster werden **ö k o n o m i s c h** (wirtschaftlich) angelegt in Abhängigkeit vom Wirkungsgrad der berücksichtigten Information. Ebenso wie der Zustandsraum sind der Wahrnehmungsraum und der Verhaltensraum (als Euklidische Räume darstellbar) in endlich viele Kategorien unterteilt. Das LS verfeinert sein Wahrnehmungs- oder Verhaltensraster, wenn die verfügbaren Kategorien zu oft keine "befriedigende" Ungleichgewichtsverringerung ermöglichen, es vergrößert, wenn ein Teil der Information bezüglich der Ungleichgewichtsänderung irrelevant ist.

Die Algorithmen der Rasterdifferenzierung bestimmen sich aus dem Vergleich der singulären aktuellen Ereignisse mit einem Stichprobenmittel bzw. einer Stichprobenvarianz der vergangenen Ereignisse.

(4) T r a n s f e r (Übertragung) von Verhaltensmodifikationen führt zu einer Mitveränderung der Verhaltensbedingtheit unter verwandten Umständen nach Maß der Verwandtschaft.

(5) Ein S y s t e m von unbedingten und bedingten S o l l - w e r t e n (deren Zahl die Dimension des Zustandsraums bestimmt) ist die Basis der Ungleichgewichtsermittlung. Zu den unbedingten Sollwerten gehören Bedingungen, die unmittelbar den Erhalt des Individuums garantieren, zu den bedingten Sollwerten, die erweitert oder aufgegeben werden können, gehören vom LS aufgenommene Sollwerte, deren Erreichen durch den Istwert die Wahrscheinlichkeit wesentlicher Ungleichgewichtssituationen als Folge der gegebenen Umweltverhältnisse verringern. Der Gleichgewichtspunkt ist der Zustand, bei dem alle Istwerte mit allen Sollwerten übereinstimmen.

Obwohl die angegebenen Algorithmen für die meisten Lernvorgänge bei geeigneter Parameterbestimmung quantitative Prognosen und damit eine experimentelle Überprüfung zulassen würden, haben wir in dieser Arbeit den Schwerpunkt auf qualitative und kategoriale Aspekte gelegt, um das Grundkonzept einigermaßen vollständig darstellen zu können. Das Modell beansprucht nicht, physiologische Vorgänge in lernenden Subjekten zu repräsentieren, aber es leistet eine ökonomische Darstellung vieler verschiedener Lernvorgänge.

Als ein wesentliches Anwendungsgebiet des hier neu konzipierten Ansatzes könnte der Bereich der Diagnose und Prognose gelten. Aufgrund der möglichen Operationalisierung komplexer Prozesse lassen sich Simulationsmodelle auf dem Computer durchrechnen, die im Vergleich mit realen Experimenten zur Parameterbestimmung oder zur Theorienmodifikation führen.

Kapitel I

ALLGEMEINES MODELL EINES LERNSYSTEMS

1.1 Grundannahmen für ein Verhaltenssystem

Wir geben zunächst einige Bezeichnungen an, die im folgenden verwendet werden. Zufällige Variable tragen grundsätzlich ein Häkchen ($\hat{X}$). Ihre Realisierungen haben meistens die gleiche oder eine ähnliche Bezeichnung, aber ohne Häkchen. Lateinische Großbuchstaben ohne Index sind Grundmengen (X), mit Index Teilmengen der betreffenden Grundmengen (X_i). Lateinische Kleinbuchstaben ohne Index bedeuten Elemente (x) einer Menge (X), mit Index sind es Teilmengen (x_i), die zugleich Elemente der zugehörigen atomaren Zerlegung sind. (vgl. Def.(1.2)). Atomare Zerlegungen werden durch deutsche Großbuchstaben mit kleinen Kreisen ($\mathcal{X}$) gekennzeichnet, ohne Kreis stellen jene einen Sigmakörper ($\mathcal{K}$) über der betreffenden Grundmenge (X) dar. Deutsche Kleinbuchstaben (μ) bezeichnen in der Regel Funktionen, mit Pfeil ($\vec{\mu}$) Vektoren, griechische Kleinbuchstaben stehen im allgemeinen für Wahrscheinlichkeitsmaße (ξ). $\mathbb{E}(\cdot)$ bezeichnet den Erwartungswert und $\text{Var}(\cdot)$ die Varianz einer zufälligen Variablen. $:=$ bedeutet die definitionsgleiche Zuordnung, $\therefore$ steht für Beweisende.

Aus der Automatentheorie entnehmen wir den Begriff des Moore-Automaten (vgl. GLUSKOW, 1963):

(1.1) Def.: Moore-Automat, Stochastischer Moore-Automat

a) $A = A(S, U, Y, f, \eta)$ heißt MOORE-AUTOMAT, wenn $S, U, Y \neq \emptyset$ und $f: S * U \rightarrow S$ sowie $\eta: S \rightarrow Y$ eindeutige Abbildungen sind.

b) Ist $T = \{\dots, 0, 1, \dots, t, \dots\}$ eine diskrete (Zeit-)indexmenge, dann heißt A STOCHASTISCHER MOORE-AUTOMAT, wenn A eine Folge von Moore-Automaten $A_t = (S, U, Y, f_t, \eta_t)$ ist, wobei $\{f_t, \tau_t, t \in T\}$ und $\{g_t, \eta_t, t \in T\}$ stochastische Prozesse mit Werten in S bzw. Y sind.

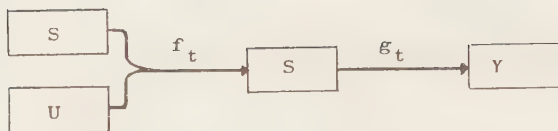


Abb. (1.1)

(1.2) Def.: Atomare Zerlegung von G

Sei $G = \mathbb{R}^m$ ($m = \text{natürlich}$). Dann heißt

$\vec{G} = \{G_1, \dots, G_n\}$ ATOMARE ZERLEGUNG von G ,

wenn gilt:

$$(1) \quad 2 \leq n < \infty$$

$$(2) \quad G_k = \bigcap_{i=1}^m I_i \quad (I_i = \text{links halboffenes Intervall} \neq \emptyset \quad (k = 1, \dots, n))$$

$$(3) \quad G_k \cap G_i = \emptyset \quad \text{für } i \neq k$$

$$(4) \quad \bigcup_{k=1}^n G_k = G$$

Mit $\vec{\mathcal{V}}(G)$ bezeichnen wir die Menge aller atomaren Zerlegungen von G .

Wir werden nun die Grundbedingungen angeben, denen das von uns betrachtete Verhaltenssystem (= VS), das als ein spezieller stochastischer Moore-Automat definiert ist, genügen muß, damit als dessen Spezialfälle die Homöostatischen Systeme (= HS) und die Lernsysteme (= LS) sinnvoll zueinander in Beziehung gesetzt werden können.

(1.3) Def.: Verhaltenssystem (=VS)

Sei $T = \{0, 1, \dots\}$, v natürlich und X, Y, Z Euklid.

Räume mit jeweils einer Dimension > 0 , sowie $S := X * Z$

und $E := S * Y$, wobei $(X, \mathcal{X})$, $(Y, \mathcal{Y})$, $(Z, \mathcal{Z})$ Maßräume

sind. Ein stochastischer Moore-Automat $A =$

$\{(S, Y, Y, f_t, g_t), t \in T\}$ heißt Verhaltenssystem mit

dem Gedächtnis v , wenn ein stochastischer Prozeß

$(\hat{E}_t, t \in T)$ mit Werten in E existiert, der durch das

Wahrscheinlichkeitsfeld $(F, \mathcal{F}, \gamma)$ mit $F := \prod_{t \in T} E(t)$,

$\mathcal{F} := \prod_{t \in T} (\mathcal{X}_{(v)} * \mathcal{Z}_{(v)} * \mathcal{Y}_{(v)})$ beschrieben wird, wobei gilt:

(1) Es ex. atomare Zerlegungen $\hat{\mathcal{X}}_0$ über $(X, \mathcal{X})$, $\hat{\mathcal{Y}}_0$ über $(Y, \mathcal{Y})$ und $\hat{\mathcal{Z}}_0$ über $(Z, \mathcal{Z})$.

(2) Für $t > 0$, $v' := \min(v, t)$ sowie $\hat{\mathcal{X}}_t := \hat{\mathcal{X}}_0 \times \hat{\mathcal{Y}}_t$ und

$\hat{\mathcal{Y}}_t = \hat{\mathcal{Y}}_0 \times \hat{\mathcal{Y}}_t$ gilt:

(a) $\hat{\mathcal{X}}_t, \hat{\mathcal{Y}}_t, \hat{\mathcal{Z}}_t$ sind Vergröberungen oder Verfeinerungen ihrer Vorgänger.

(b) Für $\hat{S}_t = s_{j_t} = (x_{j_t}, z_{j_t}) \in \hat{\mathcal{S}}_t$ ($t = t, t-1$) sei

$$[1.1] \quad \gamma(\hat{S}_t, \hat{S}_{t-1}) := \gamma(\hat{Z}_t, \hat{Z}_{t-1}) := \|z_{j_t}\| - \|\tilde{z}_{j_{t-1}}\|$$

wobei $\tilde{z}_{j_{t-1}}$ = Realisierung von $\hat{Z}_{t-1}$ bezüglich der Zerlegung $\hat{\mathcal{Z}}_t$ und

$$[1.2] \quad \|z_{j_t}\| := \inf \{ |z| : z \in z_{j_t}, || = \text{Euklid. Norm} \} \text{ ist.}$$

Dann ex. eine monoton fallende Funktion $f: \mathbb{R} \rightarrow [-1, +1]$ mit

$$[1.3] \quad \hat{c}_{t-1} = f(\hat{\gamma}_{t-1}) \text{ wobei } \hat{\gamma}_{t-1} := \gamma(\hat{S}_t, \hat{S}_{t-1}) \text{ ist.}$$

(c) Es ex. eine Abb. $f_1: \check{\mathcal{Y}}(X)^{v'} * [-1, +1]^{t-1} \rightarrow \check{\mathcal{Y}}(X)$ mit

$$[1.4a] \quad \hat{\mathcal{X}}_t = f_1(\hat{c}_{t-1}, \dots, \hat{c}_0, \hat{\mathcal{X}}_{t-1}, \dots, \hat{\mathcal{X}}_{t-v'}) \text{ die Bed. (2a) genügt.}$$

(d) Es ex. eine Abb. $f_2: \check{\mathcal{Y}}(Y)^{v'} * [-1, +1]^{t-1} \rightarrow \check{\mathcal{Y}}(Y)$ mit

$$[1.4b] \quad \hat{\mathcal{Y}}_t = f_2(\hat{c}_{t-1}, \dots, \hat{c}_0, \hat{\mathcal{Y}}_{t-1}, \dots, \hat{\mathcal{Y}}_{t-v'}) \text{ die Bed. (2a) genügt.}$$

(e) Es ex. eine Abb. $f_3: \check{\mathcal{Y}}(Z)^{v'} * [-1, +1]^{t-1} \rightarrow \check{\mathcal{Y}}(Z)$ mit

$$[1.4c] \quad \hat{\mathcal{Z}}_t = f_3(\hat{c}_{t-1}, \dots, \hat{c}_0, \hat{\mathcal{Z}}_{t-1}, \dots, \hat{\mathcal{Z}}_{t-v'}) \text{ die Bed. (2a) genügt.}$$

(3) Für $t > 0$ ist für alle $S_j \in \mathcal{S} := B(\mathcal{X} * \mathcal{Y})$, $e_i \in \mathcal{E}_{t-1}$

$$[1.5] \quad f_t(e_i, S_j) = P(\hat{S}_t = S_j | \hat{E}_{t-1} = e_i)$$

(4a) Für $t \geq 0$ ist $g_t: S * \mathcal{Y} \rightarrow R$ ein λ -stetiger Markov Kern von $(S, \mathcal{S})$ auf $(Y, \mathcal{Y})$, d.h.

$s \rightarrow g_t(s, Y_k)$ ist $\mathcal{S}$ -meßbar für alle $Y_k \in \mathcal{Y}$ und $Y_k \rightarrow g_t(s, Y_k)$ ist ein λ -stetiges Wahrscheinlichkeitsmaß auf $\mathcal{Y}$ (für alle $s \in \mathcal{S}$) (λ = Lebesguemaß auf Y)

wobei gilt:

$$[1.6a] \quad \hat{g}_t(s_{jt}, Y_k)(e_0, \dots, e_{t-1}, s_{jt}) = P(\hat{Y}_t = Y_k | \hat{E}_0 = e_0, \dots, \hat{E}_{t-1} = e_{t-1}, \hat{S}_t = s_{jt})$$

(4b) Es ex. eine Abb. $\eta = \eta_1 \circ \eta_2$, so daß für

$$\hat{g}'_{t+1}(s_j, Y_k) = \eta_1(\{\hat{g}_t(s_j, Y_k)\}_{k=1, \dots, r_t}, \hat{c}_t, \hat{E}_t = (s_{jt}, Y_{kt}))$$

mit $s_j \in \mathcal{S}_t$, $Y_k \in \mathcal{Y}_t$ gilt:

$$[1.6b] \quad \sum_{Y_k \in \mathcal{Y}_t} \hat{g}'_{t+1}(s_j, Y_k) = 1$$

Falls $K(\mathcal{S}_{t+1}) \supset K(\mathcal{S}_t)$, dann gilt für $s'_j \in \mathcal{S}_t$, mit $s'_j \in s_j \in \mathcal{S}_t$

$$\hat{g}'_{t+1}(s'_j, \cdot) = \eta_2(\hat{g}'_{t+1}(s_j, \cdot), \hat{y}_t, \dots, \hat{y}_t, \hat{c}_t):$$

$$[1.6c] \quad \hat{g}'_{t+1}(s'_j, \cdot) = \hat{g}'_{t+1}(s_j, \cdot)$$

Falls $K(\mathcal{S}_{t+1}) \subset K(\mathcal{S}_t)$, dann bestimmt sich der Wert

von $\hat{g}'_{t+1}(s'_j, \cdot)$ nach dem Satz vom iterierten Erw.wert.

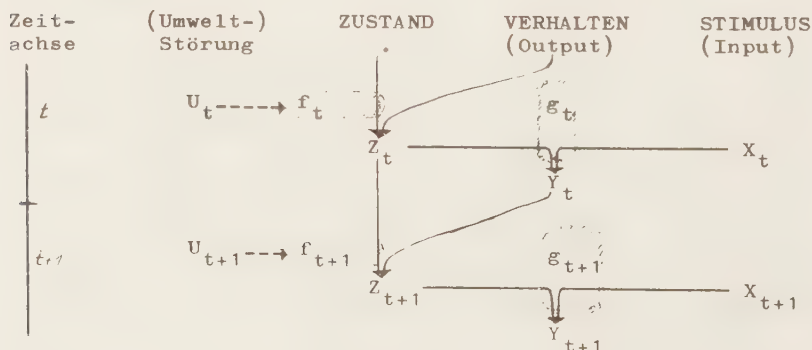


Abb. 1.2

Aus Gründen der späteren Interpretation wählen wir im folgenden spezielle Bezeichnungen für die in (1.3) eingeführten Größen: E heißt Ereignismenge, $e \in E$ Ereignis sowie $\mathcal{E}_t$ das Ereignisraster zur Zeit t . Entsprechend nennen wir S = Situationsmenge, X = Stimulusmenge, Z = Zustandsmenge und Y = Verhaltensmenge.

Als $\|z_i\|$ ($z_i \in \mathcal{Z}_t$) bezeichnen wir das Ungleichgewicht aller $z \in z_i$ und für die Situation $s_j = (x_h, z_i)$ bedeutet

$$[1.7] \quad \|s_j\| := \|z_i\|$$

das Ungleichgewicht von s_j . Das in [1.1] definierte $\hat{g}_t = \gamma(\hat{S}_t, \hat{S}_{t-1})$ heißt Ungleichgewichtsänderung zwischen Takt t und Takt $t+1$.

(1.4) Satz: Existenz des Modells

Sei v natürlich, X, Y, Z mit $\dim(.) > 0$ Euklidische Räume. $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ atomare Zerlegungen über den entsprechenden Räumen und $\hat{E}_0 = e_{i_0}$. Dann existiert für beliebig vorgegebene Abb. f, g, f_1, f_2, f_3 und Wahrscheinlichkeitsfelder $((S, \gamma, f_t), t \geq 0)$ und (Y, g, g_0) ein stochastischer Prozeß $(\hat{E}_t, t \in T)$, der den Bedingungen von (1.3) genügt und somit das Modell definiert.

Beweis: Seien die o.g. Anfangsbedingungen gegeben. Dann ist $\hat{\mathcal{Z}}_1 = \hat{\mathcal{Z}}_0$. Für $c_0 := f(\hat{g}_0) = f(\gamma(\hat{Z}_1, \hat{Z}_0))$ gibt es höchstens endlich viele Realisierungsmöglichkeiten $\hat{f}_1^{(j)}$, so daß die Menge der möglichen Ereignispaare (e_{i_0}, e_{i_1}) mit $e_{i_0} \in \hat{f}_0$, $e_{i_1} \in \hat{f}_1^{(j)}$ eine eindeutig festgelegte Menge $\hat{\mathcal{Z}}_2$ ergibt, die selbst eine atomare Zerlegung über $E_{(1)} * E_{(0)}$ ist. (vgl. Abb. (1.3)).

$g_t(\dots)$ ($t > 0$) gewinnt man aus $g_{t-1}(\dots)$ mit Hilfe der Abbildung η , wobei zunächst die Inhalte $\tilde{g}_t(\dots)$ konstruiert werden, die sich dann zu normierten Maßen auf $\mathcal{J}$ erweitern lassen. Dabei gilt für $t > 0$, wenn etwa $\mu(\check{\eta}_t) \subset \mu(\check{\eta}_{t-1})$, nach dem Satz vom iterierten Erwartungswert: $\hat{g}_{t-1}(s_j, y_k) = \sum \hat{\mu}_{t-1}(\xi \hat{\mu}_{t-1}(I_{y_k} | s_j) | \mu(\check{\eta}_t))$ für $s_j \in \check{\mathcal{S}}_t$, $s_j \subset s'_j \in \check{\mathcal{S}}_{t-1}$ und $\hat{\mu}_{t-1}(y_k | s_j) := g_{t-1}(s_j, y_k)$, wobei I_{y_k} Indikatorfunktion von y_k ist. (vgl. RICHTER, 1966; 287). Ist hingegen $\mu(\check{\eta}_t) \supset \mu(\check{\eta}_{t-1})$, dann ist die Zuordnung durch Identifizierung möglich.

Ebenso folgt für $t > 1$: $\hat{\mathcal{F}}_t = \mathcal{M}(\hat{\mathcal{F}}_{t-1}, \dots, \hat{\mathcal{F}}_{t-t'}, \hat{\mathcal{C}}_{t-1}, \dots, \hat{\mathcal{C}}_0)$ hat nur endlich viele Realisierungsmöglichkeiten $\hat{\mathcal{F}}_t^{(j)}$ und somit ist:

$$\hat{\mathcal{F}}_t := \{(e_{10}, \dots, e_{it}) \mid e_i \in \hat{\mathcal{F}}_i \mid j = f(e_{10}, \dots, e_{it-1})\}$$

eine atomare Zerlegung über $\prod_{i=0}^t E_i$.

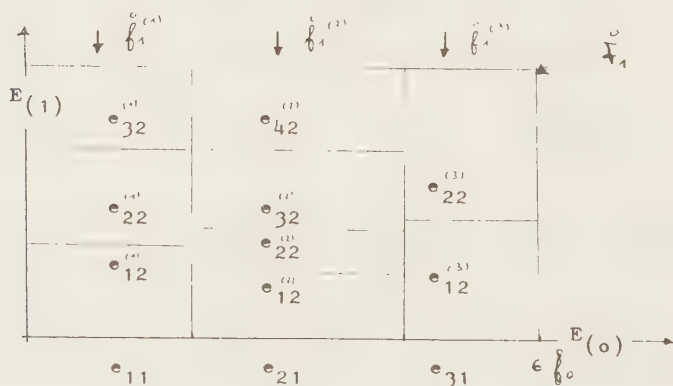


Abb. (1.3)

Nun definieren wir für $0 < t$ und $(e_{i0}, \dots, e_{it}) \in \tilde{\mathcal{F}}_t^0$ eine Mengenfunktion $p_t: \tilde{\mathcal{F}}_t^0 \rightarrow \mathbb{R}$ durch:

$$[1.8] \quad p_t(e_{i0}, \dots, e_{it}) := \prod_{\tau=0}^t [f_\tau(e_{i(\tau-1)}, s_{j\tau}) \cdot g_\tau(s_{j\tau}, y_{k\tau})]$$

und für $A \in K(\tilde{\mathcal{F}}_t^0)$

$$[1.8'] \quad p_t(A) := \sum_{a_h \in \tilde{\mathcal{F}}_t^0} p_t(A \cap a_h)$$

wobei natürlich $p_t(\emptyset) = 0$ gesetzt wird. Dies ist ein normierter Inhalt auf $K(\tilde{\mathcal{F}}_t^0)$, denn Additivität und Normiertheit folgen unmittelbar aus der Definition.

Sei nun $K(\tilde{\mathcal{F}}_t^\infty) :=$ der Körper über allen atomaren Zerlegungen über $\prod_{\tau=1}^t E_{(\tau)}$ mit einer abzählbaren Basis von links halboffenen Intervallen, deren Intervallenden dicht liegen in $\prod_{\tau=0}^t E_{(\tau)}$, mit $\nu(\tilde{\mathcal{F}}_t^0) < \nu(\tilde{\mathcal{F}}_t^\infty)$.

Dann gilt:

1) $\tilde{\mathcal{F}}_t := B(\tilde{\mathcal{F}}_t^\infty)$ ist die Menge aller Borelmengen über $\prod_{\tau=0}^t E_{(\tau)}$ mit abzählbarer Basis,

2) Es gibt eine Folge atomarer Zerlegungen $(\tilde{\mathcal{F}}_t^{(\nu)})_{\nu=1,2,\dots}$ mit

$$[1.9a] \quad \tilde{\mathcal{F}}_t^{(0)} = \tilde{\mathcal{F}}_t^0, \quad \nu(\tilde{\mathcal{F}}_t^{(\nu)}) < \nu(\tilde{\mathcal{F}}_t^{(\nu+1)}) \quad \text{und} \quad \lim_{\nu \rightarrow \infty} \tilde{\mathcal{F}}_t^{(\nu)} = \tilde{\mathcal{F}}_t^{(\infty)}.$$

In Erweiterung von [1.8] läßt sich für $(e_{i0}^{(\nu)}, \dots, e_{it}^{(\nu)}) \in \tilde{\mathcal{F}}_t^{(\nu)}$ ein Inhalt

$$[1.9b] \quad p_t^{(\nu)}(e_{i0}^{(\nu)}, \dots, e_{it}^{(\nu)}) := \prod_{\tau=0}^t [f_\tau(e_{i(\tau-1)}^{(\nu)}, s_{j\tau}^{(\nu)}) \cdot g_\tau(s_{j\tau}^{(\nu)}, y_{k\tau}^{(\nu)})]$$

definieren mit

$$[1.9c] \quad g_t^{(v)}(s_j, Y_k) := g_t(s_j, Y_k) \quad \text{für } Y_k \in \mathcal{H}$$

$$\text{ sowie } e_{i,T}^{(v)} := (s_{j,T}, y_{k,T}^{(v)}) \quad \text{ und } e_{i,T}^{(v)} \in e_{i,T} = (s_{j,T}, y_{k,T}) \in \mathcal{E}_T.$$

Für beliebige $B_t \in \mathcal{K}(\mathcal{F}_t^\infty)$ und $A_t^{(v)} = \bigcup_{i=1}^{\infty} \{ (e_{1,0}^{(v)}, \dots, e_{i,t}^{(v)}) \in \mathcal{F}_t^{(v)} : (e_{1,0}^{(v)}, \dots, e_{i,t}^{(v)}) \in B_t, t > 0 \}$ erweitern wir nun die Mengenfunktion

$$[1.9d] \quad p_\varepsilon(B_t) := \lim_{\nu \rightarrow \infty} p_t^{(v)}(A_t^{(v)}) = \\ = \lim_{\nu \rightarrow \infty} \sum_{(e_{1,0}^{(v)}, \dots, e_{i,t}^{(v)}) \in A_t^{(v)}} p_t^{(v)}(e_{1,0}^{(v)}, \dots, e_{i,t}^{(v)})$$

Der damit auf $\mathcal{K}(\mathcal{F}_t^\infty)$ gewonnene Inhalt p_t ist eindeutig, und σ -additiv, denn wegen [1.9a] gibt es zu jedem $\varepsilon > 0$, zu jedem $B_t \in \mathcal{K}(\mathcal{F}_t^\infty)$ ein $\nu_0 = \nu_0(\varepsilon, B_t)$, so daß sich B_t einbetten läßt in $A_t^{(v)} \subset B_t \subset A_t^{(v)}$ mit $p_t(A_t^{(v)} \setminus A_t^{(v)}) < \varepsilon \quad \nu \geq \nu_0$.

Daraus folgt aber die Existenz und Eindeutigkeit des Limes

[1.9d] bis auf Mengen vom Maß 0, ebenso wegen [1.8] die Normiertheit. Bleibt noch die σ -Additivität zu zeigen.

Sei $\sum_{i=1}^{\infty} B_i \subset \mathcal{F}_t$ mit $B_i \in \mathcal{K}(\mathcal{F}_t^{(v)})$.

Dann gilt für $A_i^{(v)} \rightarrow B_i \quad (A_i^{(v)} \in \mathcal{K}(\mathcal{F}_t^{(v)}))$:

$$p_t\left(\sum_{i=1}^{\infty} B_i\right) = \lim_{\nu \rightarrow \infty} p_t\left(\sum_{i=1}^{\infty} A_i^{(v)}\right) = \lim_{\nu \rightarrow \infty} p_t\left(\sum_{i=1}^n A_i^{(v)}\right) \\ \text{für geeignetes } n = n(\nu), \text{ da es in } \mathcal{F}_t^{(v)} \text{ nur endl. viele} \\ \text{Elementarereignisse gibt;} \\ = \lim_{\nu \rightarrow \infty} \sum_{i=1}^n p_t(A_i^{(v)}) = \lim_{\nu \rightarrow \infty} \sum_{i=1}^{\infty} p_t(A_i^{(v)}) = \sum_{i=1}^{\infty} p_t(B_i) \\ \text{wegen [1.9b].}$$

Ein σ -additiver, normierter Inhalt $p_t|_{\mathcal{K}(\mathcal{F}_t^\infty)}$ läßt sich aber eindeutig zu einem Maß $p_t|_{\mathcal{F}_t}$ erweitern. (vgl.

RICHTER, 1966;32). Sei nun $F := \prod_{t \in T} E_t$ und

$$\mathcal{F} := \bigcup_{t \in T} \{ A \subset F : \text{es ex. } t \in T : A = Z(A_t) \text{ und } A_t \in \mathcal{F}_t \}.$$

Dann definieren wir für $A = (E_{(0)}, \dots, A_{t_1}, E_{(t_1, 1)}, \dots, A_{t_r}, E_{(t_r, 1)}, \dots, E_t) \in \mathcal{F}$,

$$p_{t_1, \dots, t_r}(A_{t_1}, \dots, A_{t_r}) := p_t(A).$$

Dann sind die $p_{t_1, \dots, t_r}$ verträglich, wie man wegen [1.9b]

unmittelbar sieht, weil $p_t(A) = p_{t+1}(A')$ mit $A' = (A, E_{(t+1, 1)})$

für alle $t \in T$. Nach dem Satz von Kolmogoroff (vgl. RICHTER, 1966;

209), folgt daraus die Existenz eines stochastischen Prozesses

$\{\hat{E}_t, t \in T\}$ mit dem gemeinsamen Wahrscheinlichkeitsfeld

$(F, \mathcal{F}, p)$, wobei $F = \prod_{t \in T} E_t$ und $\mathcal{F} = \bigcup_{t \in T} \mathcal{F}_t$ ist und das

normierte Maß p die Eigenschaft hat. Für jede endliche Aus-

wahl $t_1, \dots, t_r \in T$ gilt: $p(Z(A_{t_1}, \dots, A_{t_r})) = p_{t_1, \dots, t_r}(A_{t_1}, \dots, A_{t_r})$

mit $A_{t_v} \in \mathcal{F}_{t_v}$ und $Z(\cdot)$ Zylindermengen über $(\cdot)$. Wegen der

angegebenen Konstruktion des Prozesses sind die Bedingungen

von (1.3) erfüllt. ./.

Bezeichnung: Im folgenden stellen wir $f_t(e_i, s_j)$ als be-

dingte Wahrscheinlichkeit $x_t(s_j | e_i)$ und $\hat{g}_t(s_j, y_k)$ durch

$\hat{\mu}_t(y_k | s_j)$ dar. Für gegebenes $\hat{S}_t = s_j$ ist $\mu_t(\cdot | s_j)$ eine

Wahrscheinlichkeitsverteilung über $(Y, \mathcal{Y})$. Wir bezeichnen

dann $\mu_t(y_k | s_j)$ als (bedingte) Verhaltenswahrscheinlichkeit.

1.2 Homöostatische Automaten und Lernsysteme

Im folgenden untersuchen wir Verhaltenssysteme, deren Verhalten (=Output) y darauf ausgerichtet ist, ein eventuell vorhandenes Ungleichgewicht zu reduzieren.

(1.6) Def.: Homöostatisches System (=HS)

Gegeben sei ein VS $A = (A_t(\cdot, v), t \in T)$ mit dem Ereignisraster $\dot{f}_t = \dot{f}$ für alle $t \in T$. Sei $\tilde{\sigma} \subset \ddot{\sigma}$ und $\tilde{\sigma}^* := \{s_j = (x_h, z_1) \in \tilde{\sigma} \mid \|z_1\| > 0\} \neq \emptyset$. A heißt HOMÖOSTATISCHES SYSTEM (=HS) für $\tilde{\sigma}$, wenn für alle $s_j \in \tilde{\sigma}^*$ gilt:

$$1) \quad v = 0$$

$$2) \quad \mu_t(\cdot | s_j) = \mu(\cdot | s_j) \quad \text{für alle } t \in T$$

$$3) \quad \text{Ist } \dot{y} = \{y_1, \dots, y_r\}, \text{ dann gilt:}$$

$$[1.10] \quad \sum_{k=1}^r \mu(y_k | s_j) * \chi_{t_i}(S1_j | s_j, y_k) = 1,$$

$$\text{wobei } S1_j := \{s_i \in \ddot{\sigma} \mid f(s_i, s_j) \leq 0\} \quad \text{ist.}$$

Eine Wahrscheinlichkeitsverteilung über $\dot{y}$, die die Bedingung [1.10] erfüllt, heißt homöostatisches Verhalten.

Die Eigenschaft, ein HS zu sein, ist also auf bestimmte Situationen bezogen. Sie stellt sowohl Bedingungen an die Umwelt - bezüglich der Reaktion auf das Verhalten - wie auch an das System selbst - bezüglich der Nutzbarmachung dieser Reaktion.

Für den Gleichgewichtszustand werden keine Bedingungen an das Verhalten des HS gestellt, weil ein System, das im Gleichgewicht bleibt, für unsere Fragestellung nicht bedeut-

Unter den möglichen Umweltreaktionen wollen wir zwei ausgewählte Fälle spezifizieren:

(1.7) Def.: Quasistabiles Situationsraster

Sei A VS mit dem Ereignisraster $\dot{f}_t = \dot{f} = \check{\sigma} * \check{\eta}$
 $\emptyset \neq \check{\sigma} \subset \check{\sigma} \quad : \quad$ Dann heißt $\check{\sigma}$ QUASISTABIL
 bezüglich $\dot{f}$, wenn für alle $s_j \in \check{\sigma}$, $y_k \in \check{\eta}$ gilt:

$$[1.11] \quad \alpha_t(.|e_i) = \alpha(.|e_i) \quad \text{für alle } t \in T, e_i = (s_j, y_k),$$

wobei $\alpha_t(.|.)$ durch [1.8] definiert ist.

(1.8) Def.: Quasideterministisches Situationsraster

Sei A VS mit dem Ereignisraster $\dot{f}_t = \dot{f} = \check{\sigma} * \check{\eta}$.
 $\emptyset \neq \check{\sigma} \subset \check{\sigma} \quad : \quad$ Dann heißt $\check{\sigma}$ QUASIDETERMINISTISCH
 bezüglich $\dot{f}$, wenn für alle $s_j \in \check{\sigma}$, $y_k \in \check{\eta}$ und
 $e_i = (s_j, y_k)$ gilt:

$$[1.12] \quad \alpha_t(.|e_i) = 0 \text{ oder } 1 \quad \text{für alle } t \in T.$$

Wir nennen ein Situationsraster $\check{\sigma}$ EINFACH, wenn es quasideterministisch und quasistabil ist.

Das im folgenden eingeführte LS soll im Gegensatz zum HS so definiert sein, daß sich im stochastischen Prozeß η_t die Verhaltenswahrscheinlichkeiten $\hat{\pi}_t$ in Abhängigkeit von den gemachten Erfahrungen ändern, mithin echte zufällige Variable sind, und zwar ändern sie sich so, daß unter möglichst schwachen Voraussetzungen ein HS approximiert wird, sofern ein solches überhaupt existiert.

(1.10) Def.: Transferkoeffizient

Sei $\check{s}_t = \{s_1, \dots, s_m\}$ ein Situationsraster eines VS.

Dann heit die Abbildung:

[1.13] $\sigma : \check{s}_t * \check{s}_t \dashrightarrow [0, 1]$ TRANSFERKOEFFIZIENT, wenn

a) $0 \leq \sigma(s_i, s_j) \leq 1$ mit $\sigma(s_i, s_i) = 1$,

b) $\sigma(s_i, s_j) = \sigma(s_j, s_i)$

c) Fr jedes $s_i \in \check{s}_t$ existiert ein $s_j \in \check{s}_t$ derart, da $\sigma(s_i, s_j) = 0$ ist. (s_j heit dann unabhngig von s_i).

(1.11) Def.: Lernsystem (=LS)

Gegeben sei ein VS $A = (A_t(v), t \in T)$ mit den Ereignisrastern $\check{b}_t = \check{s}_t * \check{y}_t$, wobei $\check{s}_t = \{s_1, \dots, s_m\}$ und $\check{y}_t = \{y_1, \dots, y_r\}$ ist. A heit LERNSYSTEM, wenn fr $\hat{E}_t = (s_{jt}, y_{kt})$ gilt: (1) f erfllt die Bedingung

[1.14] $\hat{c}_t = f(\hat{y}_t) = -Y(s_{jt+1}, s_{jt}) / Y^*(s_{jt})$

wobei $\hat{S}_{t+1} = s_{jt+1} \in \check{s}_t$ und $Y^*(s_{jt}) := \max_{s_i \in \check{s}_t} |Y(s_i, s_{jt})|$

(2) Ist $M_{t,j} = \begin{pmatrix} m_{t,1} & \dots & m_{t,r} \\ 0 & & 0 \end{pmatrix}$ mit $m_{t,k} = A_t(y_k | s_j)$,

dann erfllt η_1 die Bedingung:

[1.15] $\hat{M}_{t+1,j} = \hat{M}_{t,j} + \sigma(s_j, s_{jt}) \cdot \Theta_{t,kt}(\hat{M}_{t,j}; \hat{c}_t)$

wobei $\sigma(s_j, s_{jt})$ durch [1.13] definiert und $\Theta_{t,kt} =$

$\begin{pmatrix} \vartheta_{t,k} & 0 \\ 0 & \gamma_{t,k} \end{pmatrix}$ Lernincrement ist, d.h.

a) $-m_{t,k} < \vartheta_{t,k} < 1 - m_{t,k}$ mit $\sum_{k=1}^r \vartheta_{t,k} = 0$

b) $\vartheta_{t,kt}(\cdot, \hat{c}_t) = \max_k \vartheta_{t,k}(\cdot, \hat{c}_t)$ ist eine in $\hat{c}_t$ monoton wachsende und $\vartheta_{t,k}(\cdot, \hat{c}_t)$ ($k \neq kt$) in $\hat{c}_t$ monoton fallende Funktion mit $\vartheta_{t,k}(\cdot, 0) = 0$.

$$[1.18] \quad \hat{\gamma}_t(j,k) := \hat{\gamma}_t(j,k,jt,kt) := \\ = \sigma(j,jt) * \theta_{t,k} (s_j, x_{kt}, \hat{c}_t)$$

bezeichnen wir als LERNMODIFIKATION der Zuordnung (s_j, y_k) zur Zeit t (als Folge des Ereignisses $\hat{E}_t = (s_j, y_{kt})$).

(1.12) Def.: Nichtdifferentielles Lernsystem

Ein Lernsystem A heißt NICHTDIFFERENTIELL, wenn

$$\dot{f}_t = \dot{f} \quad \text{für alle } t \in T.$$

Schon aus der Definition des Verhaltenssystems folgt, daß ein Lernprozeß, der auf Ungleichgewichtsänderungen basieren soll, als strukturelle Basis mindestens ein einstufiges VS besitzen muß, d.h. das Ereignis des letzten Taktes $\hat{E}_t$ und $\hat{Z}_{t+1}$ muß mindestens zur Modifikation von $\vec{m}_{t,j}$ herangezogen werden. Daher müssen LS mindestens einstufige VS sein.

(1.13) Satz:

Ist für ein VS $\vec{m}_{t,j}$ auf $\check{\gamma}_t$ gegeben und tritt als Folge des Ereignisses $\hat{E}_t = (s_j, y_k)$ $\hat{c}_t = c$ auf, dann wird durch [1.15] auf $\check{\gamma}_{t+1}$ wieder ein eindeutiger Inhalt gegeben.

Beweis: Die Behauptung ergibt sich aus den Definitionen (1.11) und (1.10) sowie dem Existenzsatz (1.4) ./.

Zur konstruktiven Darstellung von Lernsystemen seien im folgenden zweckmäßige Spezifizierungen von $\sigma(...)$ und $\theta_{t,kt}(...)$ vorgenommen.

(1.14) Def.: Unabhängiges Situationsraster, Binärgitter

a) Das Situationsraster $\check{\gamma}_{(t)}^0 = \{s_{1_t}, \dots, s_{m_t}\}$ heißt UNABHÄNGIG, wenn für alle $t \in T$ gilt:

$$[1.19] \quad \sigma(s_i, s_j) = \begin{cases} 1 & \text{für } i=j \\ 0 & \text{sonst} \end{cases}$$

b) Das Situationsraster $\check{\gamma}_{(t)}^0 = \{s_{1_t}, \dots, s_{m_t}\}$ heißt BINÄRGITTER, wenn gilt für alle $t \in T$:

(1) Es gibt ein n_t mit $n_t = \text{ld}(m_t)$ und

$s_j = (a_{j,1_t}, \dots, a_{j,n_t})$ mit $a_{j,v} = 0$ oder 1 für alle $s_j \in \check{\gamma}_{(t)}^0$.

(2) Falls ν, μ zwei Komponenten sind mit der Eigenschaft, daß $a_{j,\nu} > a_{i,\nu}$ und $a_{j,\mu} < a_{i,\mu}$, dann ist $\sigma(s_i, s_j) = 0$ für $s_i, s_j \in \check{\gamma}_{(t)}^0$, andernfalls ist

$$[1.20] \quad \sigma(s_i, s_j) = \frac{\min(\|s_i\|, \|s_j\|)}{\max(\|s_i\|, \|s_j\|)}$$

wobei $\|s_i\| := \sum_{v=1_t}^{n_t} |a_{i,v}|$ ist.

(1.15) Satz:

Die durch [1.19] und [1.20] definierten Koeffizienten sind Transferkoeffizienten.

Beweis: trivial ./.

(1.16) Def.: Unabhängiges LS

Ein LS A mit $\hat{y}_t = \{y_{1t}, \dots, y_{r_t}\}$ heißt UNABHÄNGIG, wenn das Situationsraster unabhängig ist und für das in [1.15] definierte Lernincrement gilt:

$$[1.21] \quad \theta_{t,kt}(\hat{M}_{t,j}; \hat{c}_t) = \hat{c}_t^* \cdot \hat{P}_{t,j} \cdot \hat{M}_{t,j} \cdot (I - \hat{M}_{t,j}) \quad j=1_t, \dots, m_t$$

wobei I = Einheitsmatrix vom Rang r_t , $\hat{c}_t^* = 2 \cdot \hat{c}_t \cdot \min(m_{t,k}^{(j)}, 1 - m_{t,k}^{(j)})$ und $\hat{P}_{t,j} = (\hat{\rho}_t(k, kt))_{k=1_t, \dots, r_t}$ Diagonalmatrix ist mit

$$[1.22] \quad \hat{\rho}_t(kt, kt) = 1 \text{ und } \hat{\rho}_t(k, kt) = -\hat{r}_{t,kt} / \sum_{k \neq kt} \hat{r}_{t,k} \text{ für } k \neq kt$$

wobei $\hat{r}_{t,k} = m_{t,k}^{(j)} \cdot (1 - m_{t,k}^{(j)})$ ist.

(1.17) Def.: Abhängiges LS

Ein LS A mit $n_t = \text{ld } r_t$ und

$$[1.23] \quad \hat{y}_t = \{y_k = (a_{k,1_t}, \dots, a_{k,n_t}) \mid a_{k,v} = 0 \text{ oder } 1\}$$

heißt ABHÄNGIG, wenn für $p_{t,v} := \mu_t(\hat{a}_{k,v}=1 \mid s_j)$ ($v=1_t, \dots, n_t$) und

$$[1.24] \quad \mu_t(y_k \mid s_j) = \prod_{v=1_t}^{n_t} [a_{k,v} \cdot p_{t,v} + (1 - a_{k,v}) \cdot (1 - p_{t,v})] \text{ gilt:}$$

Ist $\hat{Y}_t = y_{kt} = (a_{kt,1_t}, \dots, a_{kt,n_t})$, dann folgt für die Modifikation von $P_{t,j} = (P_{t,v}^{(j)})_{v=1_t, \dots, n_t}$

$$[1.25] \quad \theta_{t,kt}(\hat{P}_{t,j}; \hat{c}_t) = c_t \cdot N(\hat{P}_{t,j}) \cdot \hat{A}_{t,kt} \cdot \hat{P}_{t,j} \cdot (I - \hat{P}_{t,j})$$

wobei I die Einheitsmatrix vom Rang $n_t = \text{ld } r_t$,

$$N(\hat{P}_{t,j}) = (2 \cdot \min(p_{t,v}^{(j)}, 1 - p_{t,v}^{(j)}))_{v=1_t, \dots, n_t} \text{ und}$$

$$\hat{A}_{t,kt} = (a_{kt,v}^{(j)})_{v=1_t, \dots, n_t} \text{ Diagonalmatrizen sind.}$$

Die expliziten Lernincremente der Verhaltensalternativen

$$y_k = (a_{k,1_t}, \dots, a_{k,n_t}) \text{ bestimmen sich dann aus [1.24].}$$

(1.18) Satz:

Für das durch [1.23] definierte Verhaltensraster gilt:
 $\sum_{k=1}^{r_t} m_t(y_k | s_j) = 1$ für alle $t \in T, j = 1_t, \dots, m_t$.

Beweis: durch Induktion: Für $n=1$, d.h. $r_t = 2$ ist die Behauptung trivial. für " $n \rightarrow n+1$ " existiert wegen [1.24] eine Darstellung für $r' = 2^n$, $r = 2^{n+1}$ (wir lassen der Einfachheit halber den Index t hier fort):

$$\begin{array}{lll} \text{1. Stelle,} & \text{n-te St.} & \text{1. St. ,... , n-te St.} \\ y_1 = (0, 0, \dots, & 0, 0); & y_{r'+1} = (0, 0, \dots, & 0, 1) \\ y_2 = (1, 0, \dots, & 0, 0); & y_{r'+2} = (1, 0, \dots, & 0, 1) \\ \dots & & \dots \\ y_{r'} = (1, 1, \dots, & 1, 0); & y_r = (1, 1, \dots, & 1, 1) \end{array}$$

Daraus folgt aber:

$$\begin{aligned} \sum_{k=1}^{r'} m_t(y_k | s_j) &= \sum_{k=1}^{r'} \prod_{v=1}^{n+1} [(1-p_{t,v}) \cdot (1-a_{k,v}) + p_{t,v} \cdot a_{k,v}] \\ &= p_{t,n+1} \cdot \sum_{k=1}^{r'} \prod_{v=1}^n [(1-p_{t,v})(1-a_{k,v}) + p_{t,v} a_{k,v}] + \\ &+ (1-p_{t,n+1}) \cdot \sum_{k=1}^{r'} \prod_{v=1}^n [\dots] = p_{t,n+1} + (1-p_{t,n+1}) = 1 \\ &\quad \text{nach Induktionsannahme} \quad ./. \end{aligned}$$

(1.20) Satz:

Die in [1.24] und [1.25] definierten $\theta_{t,kt}(\cdot; \cdot)$ erfüllen die Bedingungen von (1.41) an ein Lern-increment.

Beweis: 1) Sei $\theta_{t,kt}(\cdot)$ definiert durch [1.24].

Dann ist die Bed. (a) erfüllt, da $|c_t^*| \leq 1$ und $|\vartheta(k, kt)| < 1$,

wie man wegen $m_t(y_{kt}|s_j) * (1 - m_t(y_{kt}|s_j)) \leq$

$\sum_{k \neq kt} m_t(y_k|s_j) * (1 - m_t(y_k|s_j))$ unmittelbar schließt.

$$\begin{aligned} \sum_{k=1}^r \vartheta_{t,k} &= c_t * [m_t(y_{kt}|s_j) * (1 - m_t(y_{kt}|s_j)) - \\ &- m_t(y_{kt}|\cdot) * (1 - m_t(y_{kt}|\cdot)) * \frac{\sum_{k \neq kt} m_t(y_k|\cdot) * (1 - m_t(y_k|\cdot))}{\sum_{k \neq kt} m_t(y_k|\cdot) * (1 - m_t(y_k|\cdot))}] \\ &= 0. \end{aligned}$$

$$(c) \quad \left| \sum_{k \neq kt} \vartheta_{t,k} \right| = |\vartheta_{t,kt}| \quad \text{für } c_t \neq 0 \quad \text{und wegen [1.22]}$$

sind alle $\vartheta_{t,k} \quad (k \neq kt)$ von gleichem Vorzeichen. Der Rest

folgt unmittelbar aus [1.24].

2) Sei $\theta_{t,kt}(\cdot)$ durch [1.25] definiert.

(a) Es genügt zu zeigen, daß

$$\sum_{k=1}^r m_{t,k} = 1 \quad \text{ist. Beweis durch Induktion: (Sei } p_{t,v} := \mu_t(a_v) \text{ und } \vartheta_{t,v} := \text{Lernincrement von } p_{t,v} \text{)}$$

$$\text{"n=1": } \sum_{k=1}^2 \prod_{v=1}^1 [(1 - p_{t,v} + \vartheta_{t,v}) * (1 - a_{k,v}) + (p_{t,v} - \vartheta_{t,v}) * a_{k,v}]$$

$$= (1 - p_{t,v} + \vartheta_{t,v}) * \underbrace{(1 - a_{1,1})}_{=1} + (p_{t,v} - \vartheta_{t,v}) * \underbrace{a_{2,1}}_{=1} = 1.$$

"n $\rightarrow$ n+1": Darstellung der y_k wie in Satz (1.18). $r = 2^{n+1}$

und $r' = 2^n$:

$$\begin{aligned}
 \sum_{k=1}^r m_{t+1,k} &= \sum_{k=1}^r \prod_{v=1}^{n+1} [(1-p_{t,v} + \mathcal{I}_{t,v}) \cdot (1-a_{k,v}) + \\
 &\quad + (p_{t,v} - \mathcal{I}_{t,v}) \cdot a_{k,v}] = \\
 &= (1-p_{t,n+1} + \mathcal{I}_{t,n+1}) * \sum_{k=1}^{r'} \prod_{v=1}^n [(1-p_{t,v} + \mathcal{I}_{t,v}) * \\
 &\quad * (1-a_{k,v}) + (p_{t,v} - \mathcal{I}_{t,v}) \cdot a_{k,v}] + \\
 &\quad + (p_{t,n+1} - \mathcal{I}_{t,n+1}) * \sum_{k=1}^{r'} \prod_{v=1}^n [...] \\
 &= 1 \\
 &\quad \uparrow \\
 &\text{nach Induktionsannahme .}
 \end{aligned}$$

(b) Falls $\hat{f}_t \geq 0$, verändert sich jede der unabhängigen Komponenten so, daß y_{kt} wahrscheinlicher wird. Folglich wird ein y_k im gleichen Maße zunehmen, wie es in den Komponenten mit y_{kt} übereinstimmt, in den andern wird es abnehmen. Das zu y_{kt} duale $y_{kt}^* = ((1-a_{kt,1}), \dots, (1-a_{kt,n}))$ nimmt genauso stark ab, wie y_{kt} zunimmt, folglich ist die Bedingung (b) erfüllt. Analog geht der Schluß, wenn $\hat{f}_t < 0$ ist. Der Rest ist trivial. ./.

(1.21) Satz:

- a) Der durch [1.26] definierte Modifikationsprozeß $\{\hat{P}_t(j,k), t \geq 0\}$ mit $s_j \in \check{\mathcal{S}}_0$ und $y_k \in \check{\mathcal{Y}}_0$ ist ein stochastischer Prozeß.
- b) Sei $s_j \in \check{\mathcal{S}}_0$, $Y_h \in \mathcal{V}(\check{\mathcal{Y}}_0)$ mit $Y_h \neq \emptyset$. Dann ist auch der durch [1.28] definierte Modifikationsprozeß bezüglich (s_j, Y_h) ein stochastischer Prozeß.

Beweis: a) Zu zeigen ist, daß durch eine Version von $\text{pr}(\hat{\mu}_t(\dots) \mid \mathcal{I}_t)$, die den Randbedingungen innerhalb der Klammer von [1.26b] genügt, eine Wahrscheinlichkeitsverteilung definiert ist. Das folgt aber unmittelbar aus Satz (1.4), wo unter Berücksichtigung der angegebenen Nebenbedingungen eine Wahrscheinlichkeitsverteilung $\text{pr}(\cdot)$ konstruktiv definiert wird. Lediglich bleibt zu beachten, daß $\hat{P}_t(j,k) \stackrel{\text{f.ä.}}{=} 0$ gesetzt wird, wenn sich die atomaren Zerlegungen $\check{\mathcal{S}}_t, \check{\mathcal{Y}}_t$ so verändern, daß s_j bzw. y_k keine Elemente mehr von ihnen sind.

b) Da die $y_k \in \check{\mathcal{Y}}_0$ paarweise disjunkte Mengen sind, ist die Bedingung $0 \leq \pi(\cdot) \leq 1$ trivialerweise erfüllt.

Auch die Normiertheit folgt unmittelbar, wenn man

$$\check{\mathcal{Y}}^* = \{Y_h, Y \setminus Y_h\} \text{ als Vergrößerung von } \check{\mathcal{Y}}_0 \text{ auffaßt.}$$

Die übrigen Bedingungen ergeben sich aus a). ./.

Nützlich sind folgende allgemeinen Aussagen über den Modifikationsprozeß:

(1.22) Satz:

Sei $\{\hat{\mathcal{P}}_t(j, k), t \geq 0\}$ der durch [1.26] definierte Modifikationsprozeß eines LS. Dann gilt:

a) $0 \leq \mathbb{E}(\hat{\mathcal{P}}_t(j, k)) \leq 1$

b) $0 \leq \text{VAR}(\hat{\mathcal{P}}_t(j, k)) \leq 1/4$

c) Für den bedingten Erwartungswert gilt:

[1.29] $\mathbb{E}(\{\hat{\mathcal{P}}_{t+1}(\cdot) \mid \hat{E}_T(\tau \leq t), \hat{S}_t = s_{jt}\}) =$

$$= \hat{\mathcal{P}}_t + \mathbb{E}(\{\hat{\eta}_t(j, k) \mid \hat{E}_T(\tau \leq t), \hat{S}_t = s_{jt}\})$$

wobei $\hat{\eta}_t(j, k)$ die in [1.18] definierte Lernmodifikation ist.

d) Ist θ ein unabhängiges LS mit $\hat{S}_t = s_{jt}$, dann gilt für den bedingten Erwartungswert der Lernmodifikation (wobei $\hat{\eta}_t = \eta$ für alle $t \in T$):

[1.30] $\mathbb{E}(\{\hat{\eta}_t(j, k) \mid \hat{E}_T(\tau \leq t), \hat{S}_t = s_{jt}\}) = \theta(j, j_t) \cdot \hat{r}_{t, k}^{(j)}$

$$+ 2 \sum_{j \neq k} \hat{m}_{t, kt}^{(j)} \cdot \min(\hat{m}_{t, kt}^{(j)}, 1 - \hat{m}_{t, kt}^{(j)}) \cdot \hat{\mathcal{P}}_t(k, kt) \cdot \mathbb{E}(\hat{c}_t \mid s_{jt}, y_{kt})$$

wobei $m_{t, kt}^{(j)} = \mu_t(y_{kt} \mid s_j)$ und $r_{t, k}^{(j)} = m_{t, k}^{(j)} \cdot (1 - m_{t, k}^{(j)})$ sowie

$$\mathbb{E}(\hat{c}_t \mid s_{jt}, y_{kt}) = -1/\gamma^*(s_{jt}) \cdot \sum_{s_i \in \mathcal{S}} \gamma(s_i, s_{jt}) \cdot \pi_t(s_i \mid s_{jt}, y_{kt})$$

e) Bei einem unabhängigen (d.h. nichtdifferentiellen) LS gilt für alle Modifikationsprozesse $\{\hat{\mathcal{P}}_t(j, k), t \geq 0\}$:

$$0 < \mathbb{E}\hat{\mathcal{P}}_t(j, k) < 1$$

f) Ist $Y = \sum_{h=1}^n Y_h$ mit $Y_h \in K(\mathcal{Y}_t)$, dann gilt für die zugehörigen Modifikationsprozesse eines LS:

$$\sum_{h=1}^n \hat{\mathcal{P}}_{h_t}(j) = 1 \quad \text{und} \quad \sum_{h=1}^n \hat{\eta}_{h_t}(j) = 0$$

Beweis: a) trivial wegen [1.27].

b) Da $\forall \hat{\mathcal{P}}_t(.) \leq \xi \hat{\mathcal{P}}_t(.) * (1 - \xi \hat{\mathcal{P}}_t(.))$ (z.B. nach RICHTER, 1966; 249), folgt die Behauptung aus a).

[1.30]

c) Die Behauptung $\forall$ folgt unmittelbar aus (1.3) und der Definition des LS in (1.11).

d) Die Behauptung folgt aus der Definition des unabh. LS, denn nach (1.16) ist für $\hat{S}_t = s_{jt}$ und $\hat{Y}_t = y_{kt}$ für beliebige y_k das Lernincrement

$$\hat{\mathcal{J}}_{t,k}(s_j, y_{kt}) = \hat{c}_t \cdot \hat{r}_t(y_k | s_j) \cdot \hat{\xi}_t(k, kt) \cdot 2 \cdot \min(m_{t,k}, 1 - m_{t,k})$$

Da $\hat{\mathcal{J}}_t(s_{jt}, y_{kt}) = \mathcal{J}(s_i, s_{jt})$ für $\hat{S}_{t+1} = s_i \in \mathcal{S}$ und andererseits $\hat{S}_{t+1} = s_i$ mit der bedingten Wahrscheinlichkeit

$\alpha_t(s_i | s_{jt}, y_{kt})$ bei unabhängigen LS, falls $\hat{S}_t = s_{jt}$ und $\hat{Y}_t = y_{kt}$, folgt:

$$\begin{aligned} \xi(\hat{c}_t | s_{jt}, y_{kt}) &= -1/\mathcal{J}^*(s_{jt}) * \int_{\mathcal{S}} \mathcal{J}(s, s_{jt}) \cdot d\alpha_t(s | s_{jt}, y_{kt}) \\ &= -1/\mathcal{J}^*(s_{jt}) * \sum_{s_i \in \mathcal{S}} \mathcal{J}(s_i, s_{jt}) \cdot \alpha_t(s_i | s_{jt}, y_{kt}). \end{aligned}$$

e) Nach der Definition des Lernincrements und des LS folgt, daß $\hat{\mathcal{P}}_t(j, k)$ π -fast sicher im offenen Intervall (0,1) liegen muß. Das impliziert die Behauptung.

f) Nach Definition der Lernmodifikation muß die Summe der Modifikationen gleich Null und daher die Summe der $\hat{\mathcal{P}}_t(j, k)$ gleich 1 sein. Daraus folgt aber die Behauptung für jede Vergrößerung von $\hat{\mathcal{J}}_t$. //

In Kapitel 2 werden wir nun zeigen, daß die in [1.26] bzw. [1.28] definierten Prozesse wünschenswerte Konvergenzeigenschaften besitzen, die die Brauchbarkeit der LS-Algorithmen

zur Approximation an ein homöostatisches Verhalten unabhängig von der Ausgangssituation nachweisen. Gemäß den Definitionen der Maßtheorie führen wir die starke und die schwache Konvergenz sowie die Konvergenz im quadratischen Mittel ein (vgl. z.B. RICHTER, 1966; Kap.7).

(1.23) Def.: π -fast sichere (starke) Konvergenz

Sei $\hat{\mathcal{P}}_t(.)$ der durch [1.26] bzw. [1.28] definierte Modifikationsprozeß. $\hat{\mathcal{P}}_t(.)$ KONVERGIERT STARK (= π -fast sicher), wenn ein $\hat{\mathcal{P}}_\infty$ existiert mit:

$$[1.31] \quad \pi \left(\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t(.) = \hat{\mathcal{P}}_\infty \right) = 1$$

(1.24) Def.: Stochastische (schwache) Konvergenz

Sei $\hat{\mathcal{P}}_t(.)$ wie vorher definiert. $\hat{\mathcal{P}}_t(.)$ KONVERGIERT STOCHASTISCH (schwach), wenn ein $\hat{\mathcal{P}}_\infty$ existiert mit

$$[1.32] \quad \pi \lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t(.) = \hat{\mathcal{P}}_\infty$$

(1.25) Def.: Konvergenz im Quadratmittel

$\hat{\mathcal{P}}_t(.)$ (wie vorher definiert) KONVERGIERT IM QUADRATMITTEL, wenn es ein $\hat{\mathcal{P}}_\infty$ gibt mit:

$$[1.33] \quad \lim_{t \rightarrow \infty} \mathbb{E} \left(\left(\hat{\mathcal{P}}_t(.) - \hat{\mathcal{P}}_\infty \right)^2 \right) = 0$$

Für den Nachweis der in dem nächsten Kapitel behaupteten Zusammenhänge sind einige bekannte Zusammenhänge über Konvergenzen von Nutzen, die der Einfachheit halber im nachfolgenden Satz zusammengefaßt werden, spezialisiert auf den von uns betrachteten Modifikationsprozeß.

(1.26) Satz:

Sei $\hat{\mathcal{P}}_t(.)$ der durch [1.26] bzw. [1.28] definierte Modifikationsprozeß. Dann gilt:

- a) $\hat{\mathcal{P}}_t(.)$ konvergiert π -fast sicher $\rightarrow \hat{\mathcal{P}}_t(.)$ konvergiert stochastisch.
- b) $\hat{\mathcal{P}}_t(.)$ konvergiert stochastisch $\leftrightarrow \hat{\mathcal{P}}_t(.)$ konvergiert im Quadratmittel.
- c) $\hat{\mathcal{P}}_t(.)$ konvergiert im Quadratmittel gegen 0 (bzw. 1) $\leftrightarrow \mathbb{E}\hat{\mathcal{P}}_t(.)$ konvergiert gegen 0 (bzw. 1).
- d) $\hat{\mathcal{P}}_t(.)$ ist π -fast sicher monoton nicht fallend $\rightarrow \hat{\mathcal{P}}_t(.)$ konvergiert π -fast sicher.
- e) Gilt $\pi\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t(.) = c_1$ und $\pi(\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t(.) = c_2) = 1$, dann folgt: $c_1 = c_2$.
- f) Falls $\lim_{t \rightarrow \infty} \mathbb{E}\hat{\mathcal{P}}_t(.) = 0$ oder 1, dann gilt auch:
 $\lim_{t \rightarrow \infty} \forall \omega \hat{\mathcal{P}}_t(.) = 0$.

Beweis: a) folgt aus der Maßtheorie. (vgl. z.B. RICHTER, 1966; 391).

b) gilt nach RICHTER(1966, 391), weil alle $\hat{\mathcal{P}}_t(.) \leq 1$, also gleichmäßig beschränkt.

c) Da $0 \leq \hat{\mathcal{P}}_t(.) \leq 1$, folgt für $c = 0$ oder 1 :
 $\lim_{t \rightarrow \infty} \mathbb{E}((\hat{\mathcal{P}}_t - c)^2) = 0$ genau dann, wenn $\lim_{t \rightarrow \infty} \mathbb{E}(\hat{\mathcal{P}}_t - c) = 0$.

d) folgt nach dem Satz v.d.maj.Kgz (vgl. z.B. RICHTER, 1966; 394).

e) folgt, da jeder stochastischer Grenzwert eind. ist.

f) Da $\forall \omega \hat{\mathcal{P}}_t(.) \leq \mathbb{E}\hat{\mathcal{P}}_t(.) * (1 - \mathbb{E}\hat{\mathcal{P}}_t(.))$, folgt die Behauptung.

./.

Kapitel II

GRENZWERTSÄTZE ÜBER LERNPROZESSE

2.1 Konvergenzkriterien für unabhängige Lernsysteme

Die folgenden Aussagen in § 2.1 und § 2.2 beziehen sich alle auf nichtdifferentielle Lernsysteme. Jene sind immer gemeint, wenn nur von unabhängigen oder abhängigen LS gesprochen wird.

Wesentliches Hilfsmittel bei der Untersuchung des Konvergenzverhaltens ist die Berechnung der "erwarteten Ungleichgewichtsänderung" $\xi(y_t | s_{jt}, y_{kt})$, die sich nach [1,30] angeben läßt durch:

$$[2.1] \quad \xi(y_t | s_{jt}, y_{kt}) = \sum_{s_i \in \tilde{\mathcal{S}}} y(s_i, s_{jt}) \cdot x_t(s_i | s_{jt}, y_{kt})$$

falls $\hat{E}_t = (s_{jt}, y_{kt})$ ist.

(2.1) Def.: Wesentlicher Rasterausschnitt

Sei das Situationsraster eines VS A . Dann heißt $\tilde{\mathcal{S}}$ mit $\emptyset \neq \tilde{\mathcal{S}} \subset \mathcal{S}$ WESENTLICHER RASTERAUSSCHNITT, wenn für alle $s_j \in \tilde{\mathcal{S}}$ gilt:

- (1) $s_j = (x_j, z_j)$ mit $\|z_j\| > 0$
- (2) Es eine abz.unendliche Teilmenge $T^* \subset T$, so daß

$$[2.1a] \quad x_{t_1}(s_j | s_i, \cdot) = 1$$

für alle $s_i \in \tilde{\mathcal{S}}$, für alle $t_1 \in T^*$.

Die in (2.1) angegebene Bedingung garantiert, daß die zum Lernen notwendigen Situationen für das LS auch tatsächlich hinreichend oft auftreten.

(2.2) Def.: Approximation eines HS

Sei $A = (A_t, t \in T)$ ein VS über dem Ereignisraum $\hat{\mathcal{F}}_t = \hat{\mathcal{F}} = \check{\mathcal{V}} * \check{\mathcal{Y}}$ mit den Modifikationsprozessen $\{\hat{\mathcal{P}}_t(j, k), t \in T\}$ für $(s_j, y_k) \in \hat{\mathcal{F}}$. Sei $\tilde{\mathcal{V}} \subset \check{\mathcal{V}}$ mit

$$[2.2a] \quad \tilde{\mathcal{V}} = \{s_j = (x_h, z_i) \in \check{\mathcal{V}} \mid \|z_i\| > 0\}$$

$$[2.2b] \quad Y1^*(j) := \{y_k \in \check{\mathcal{Y}} \mid \lim_{t \rightarrow \infty} \alpha_t(S1_j \mid s_j, y_k) = 1\}$$

wobei

$$[2.2c] \quad S1_j := \{s_i \in \tilde{\mathcal{V}} \mid \gamma(s_i, s_j) < 0\}$$

Dann APPROXIMIERT A ein HS auf $\tilde{\mathcal{V}}$, wenn für alle $s_j \in \tilde{\mathcal{V}}$ gilt:

$$[2.3] \quad \pi \left(\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_{1_t}(j) = 1 \right) = 1$$

wobei $\hat{\mathcal{P}}_{1_t}(j)$ der Modifikationsprozess bezüglich des Ereignisses $(s_j, Y1^*(j))$ ist.

Im folgenden sei zunächst ein spezieller Fall herausgegriffen, nämlich das Vorhandensein einfacher Situationen (vgl. Def. (1.7) und (1.8)). Hier kann der Nachweis der Approximation noch mit relativ elementaren Mitteln geführt werden.

(2.3) Satz:

Gegeben sei ein unabhängiges LS A mit dem Ereignis-
raster $\hat{\mathcal{F}} = \hat{\mathcal{T}} * \hat{\mathcal{Y}}$ und beliebigen Anfangswerten $\hat{\mu}_0(y_k | s_j)$
 $= \alpha_{jk}$ für $(s_j, y_k) \in \hat{\mathcal{F}}$. $\tilde{\mathcal{T}}$ sei einfach und we-
sentlicher Rasterausschnitt, wobei $\emptyset \neq \tilde{\mathcal{T}} \subset \hat{\mathcal{T}}$ gilt.
Ferner gibt es ein VS auf $\hat{\mathcal{F}}$, das für $\tilde{\mathcal{T}}$ ein HS ist.
Beh.: A approximiert auf $\tilde{\mathcal{T}}$ ein HS.

Beweis: Seien $\tilde{\mathcal{T}}, Y1*_j$ und $S1_j$ wie in [2.2] definiert.

Dann gilt wegen der Einfachheit für ein bel. $s_j \in \tilde{\mathcal{T}}^*$:

$Y1*_j = \{y_k \in \hat{\mathcal{Y}} \mid \alpha(S1_j | s_j, y_k) = 1\}$ und für $Y2_j := Y \setminus Y1*$

$Y2_j = \{y_k \in \hat{\mathcal{Y}} \mid \alpha(S2_j | s_j, y_k) = 1\}$, wobei

$$[2.4] \quad S2_j := \{s_i \in \tilde{\mathcal{T}} \mid \gamma(s_i, s_j) \geq 0\}$$

Da nach Voraussetzung ein HS auf $\tilde{\mathcal{T}}$ existiert, folgt $Y1*_j \neq \emptyset$.

Ist nun $Y2_j = \emptyset$, dann wird die Behauptung trivial.

OEdA sei daher $Y2 \neq \emptyset$. Sind $\hat{\mathcal{P}}_{1t}(j)$ und $\hat{\mathcal{P}}_{2t}(j)$ die zu $Y1*_j$ bzw. $Y2_j$ gehörigen Modifikationsprozesse, dann ist zu zeigen; daß [2.3] gilt und demzufolge auch:

$$\pi \left(\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_{2t}(j) = 0 \right) = 1$$

Zur Vereinfachung der Darstellung wählen wir folgende Notation:

$$\hat{p}_{t,k} := \hat{\mu}_t(y_k | s_{jt})$$

$$\hat{q}_{t,k} := 1 - \hat{p}_{t,k}, \quad \min_{t,k} = \min(\hat{p}_{t,k}, \hat{q}_{t,k})$$

$$\hat{r}_{t,k} := \hat{p}_{t,k} * \hat{q}_{t,k}$$

für das im folgenden fest angenommene $s_{jt} \in \tilde{\mathcal{T}}$. Für $j \neq jt$ ändert sich nach (1.1) keine der Verhaltensweisen auf s_{jt} , so daß es genügt, die Teilfolge $T^* := \{t \in T \mid \hat{S}_t = s_{jt}\}$

zu betrachten, die wegen $s_{jt} \in \tilde{\mathcal{T}}$ ^{p -fast sicher} abzählbar unendlich ist. Ferner sei

$$\hat{\eta}_t(k, kt)$$

die nach [1.18] definierte Lernmodifikation für $\hat{S}_t = s_{jt}$. Dann gelten wegen der Einfachheit von $\tilde{\mathcal{T}}$ für die Lernmodifikationen folgende Abschätzungen (π -fast sicher):

$$\hat{\eta}_t(k, kt) = \begin{cases} \geq \min_{t,k} \hat{c}_t \cdot \hat{r}_{t,k} > 0 & \text{für } y_k = y_{kt} \in Y1^* \\ \leq \min_{t,k} \hat{c}_t \cdot \hat{\beta}(k, kt) \cdot \hat{r}_{t,k} < 0 & \text{" } y_k \neq y_{kt} \in Y1^* \\ \leq 0 & \text{" } y_k = y_{kt} \in Y2 \\ \geq 0 & \text{" } y_k \neq y_{kt} \in Y2 \end{cases}$$

wobei $\hat{\beta}(k, kt)$ der durch [1.23] definierte Normierungsfaktor ist. Folglich ergibt sich für

$$\hat{\eta}_t^1 = \begin{cases} \hat{r}_{t,kt} + \hat{\beta}(k, kt) \sum_{\substack{k \neq kt \\ y_k \in Y1^*}} \hat{r}_{t,k} > 0 & \text{für } Y_t = y_{kt} \in Y1^* \\ \geq 0 & \text{für } Y_t = y_{kt} \in Y2 \end{cases}$$

wobei $c_{kt} := -f(s_i, s_j) / f^*(s_j)$ mit $\pi(s_i | s_j, y_{kt}) = 1$

d.h. $\hat{\mathcal{P}}_{1t} \geq \hat{\mathcal{P}}_{1t-1}$ π -fast sicher für alle $t \in T$

Wegen der Beschränktheit der $\hat{\mathcal{P}}_{1t}$ folgt die Konvergenz (s. (1.26))

Annahme: es gibt ein $\mathfrak{A}_1$ mit $0 < \mathfrak{A}_1 < 1$, so daß

$$\hat{\mathcal{P}}_{2t} \geq \mathfrak{A}_1 \quad \pi\text{-fast sicher für alle } t \in T.$$

Setze: $\beta := 1 - \hat{\mathcal{P}}_{20} > 0$. Dann folgt: $\hat{q}_{t,k} > \beta > 0$

π -fast sicher für alle $y_k \in Y2$, für alle $t \in T$.

Daraus schätzen wir ab:

$$\sum_{y_k \in Y_2} \hat{r}_{t,k} > \beta \cdot \min_{k \in K} \sum_{y_k \in Y_2} \hat{p}_{t,k} \geq \min_{k \in K} \beta \cdot \hat{r}_1 > 0 \quad \pi\text{-fast sicher.}$$

$$\text{Weiter ist } \sum_{k \neq kt} \hat{r}_{t,k} < \hat{q}_{t,kt} \cdot \max_{k \neq kt} \hat{q}_{t,k} < \hat{q}_{t,kt}$$

d.h.

$$|\eta_2(t)(kt)| > \frac{c^* \cdot \hat{p}_{t,kt} \cdot \hat{q}_{t,kt} \cdot \hat{r}_1 \cdot \beta \cdot \min_{k \in K} \hat{q}_{t,k}}{\hat{q}_{t,kt}} =: \hat{r}_2 \cdot \hat{p}_{t,kt} \quad \pi\text{-fast sicher}$$

mit $\hat{r}_2 > 0$, wenn man $c^* := \min_{y_{kt} \in Y_1^*} c_{kt} \quad (> 0)$ setzt. Da aber

$\tilde{\gamma}$ wesentlich ist, können wir in Verallgemeinerung des Lemmas 2 von BOREL-CANTELLI (vgl. RICHTER, 1966; 400) einen Satz von DOOB verwenden, der das hinreichend häufige Auftreten von $\hat{Y}_t \in Y_1^*$ garantiert.

Da $\hat{\mu}_t(Y_1^* | s_j) \geq \hat{\mu}_0(Y_1^* | s_j) =: \alpha > 0$, gilt wegen

$$0 < r_1 := |Y_1^*| < \infty \quad \text{für ein } y_{k_0} \quad \text{mit}$$

$$\hat{\mu}_{t_0}(y_{k_0} | s_j) := \max_{y_k \in Y_1^*} \hat{\mu}_{t_0}(y_k | s_j) \geq \alpha / r_1$$

Sei nun für $t > 0$, natürlich, und $T^* := \{t > 0, \hat{S}_t = s_j\} = \{t_i\}_{i=1,2,\dots}$

(wobei die t_i jeweils von der Vergangenheit abhängen können):

$$M_h := M_h(s_j) := \{ \hat{S}_{t_h} = s_j, \hat{Y}_{t_h} = y_{k_0} \}$$

und

$$p_h(M_h | M_1, \dots, M_{h-1}) = \hat{\mu}_{t_h}(\hat{Y}_{t_h} = y_{k_0} | \hat{E}_{t_i} = (s_j, y_{k_0}) \quad (i=1, \dots, h))$$

$$\geq \alpha / r_1 \quad \text{für alle } t_h \in T^*$$

Dann ist

$$\lim_{h \rightarrow \infty} \sum_{i=1}^h P_{t_i}(y_{k_0} | s_j) \geq \lim_{h \rightarrow \infty} h \cdot \alpha / r_1 = \infty$$

Nach DOOB (1965;324) tritt dann $\hat{Y}_t = y_{k_0}$ mit Wahrscheinlichkeit 1 unendlich oft ein und daher ist

$-\hat{\eta}_t^2(k_t) > \beta_2 \cdot \alpha / r_1 > 0$ unabhängig von t $\bar{\pi}$ -fast sicher unendlich oft. Dies ist aber ein Widerspruch zur

Tatsache, daß $\hat{\beta}_t^2 \geq 0$ für alle $t \in T$. Folglich gilt:

$$\bar{\pi}(\lim \hat{\beta}_t^2 = 0) = 1.$$

Annahme 2: $\bar{\pi}(\hat{\beta}_t^2 > \beta_1) \geq \beta_1$, aber es ex. ein $\epsilon > 0$:

zu jedem $n \in \mathbb{N}$ gibt es ein $T^*(n) \subset T$, abz. unendlich,

so daß $\bar{\pi}(\hat{\beta}_t^2 < 1/n) \geq \epsilon$.

Dann folgt wegen $\hat{\beta}_t^2 \leq \hat{\beta}_{t-1}^2$ $\bar{\pi}$ -fast sicher: Es gibt

$\bar{\pi}$ -fast sicher ein $t_n \in T^*(n)$, so daß $\hat{\beta}_{t_n}^2 < 1/n$ für

alle $t \geq t_n$, d.h. aber $\hat{\beta}_t^2$ konvergiert $\bar{\pi}$ -fast sicher gegen 0. Folglich gilt die Behauptung. ./.

Bemerkung: Man beachte, daß $\hat{A}$ keineswegs ein vorgegebenes HS approximieren muß. Dies geht schon deshalb nicht, weil für Situationen im Gleichgewichtszustand (also für $|z_i| = 0$) keine Bedingungen an ein HS gestellt werden, und die bedingten Verhaltenswahrscheinlichkeiten dort in beliebiger Weise geändert werden können.

Die Brauchbarkeit des Lernalgorithmus zur Approximation von HS ist damit für einen ausgezeichneten Fall von Umweltbedingungen gezeigt worden, nämlich für einfache Situationen. Die Approximation ist jedoch auch noch unter wesentlich allgemeineren Bedingungen möglich, wie im folgenden gezeigt werden wird.

(2.4) Satz: Approximationssatz für unabh. LS

$\tilde{\gamma}$ sei das Raster eines unabh. LS und $\emptyset \neq \tilde{\gamma} \subset \tilde{\gamma}$ sei wesentlicher Rasterausschnitt. Für jedes $s_j \in \tilde{\gamma}$ ex. eine Zerlegung

$$[2.5a] \quad \tilde{\gamma} = Y1_t + Y2_t \quad \text{mit} \quad E(\tilde{\gamma}_t | s_j, Y2_t) \geq 0$$

für t hinr. groß und der Eigenschaft, daß $Y1_t$ schließlich konstant $= Y1 \neq \emptyset$ ist.

$$[2.5b] \quad \sum_{t=1}^{\infty} \left[\sum_{y_k \in Y1_t} \alpha_t(S2_j | s_j, y_k) + \sum_{y_k \in Y2_t} \alpha_t(S1_j | s_j, y_k) \right] < \infty$$

wobei $S1_j := \{s_1 \in \tilde{\gamma} \mid \gamma(s_1, s_j) < 0\}$ und $S2_j := \tilde{\gamma} \setminus S1_j$ ist.

Beh.: A approximiert auf $\tilde{\gamma}$ ein HS.

Beweis: Sei $s_j \in \tilde{\gamma}$ (im folgenden fest). Dann gibt es ein $c' > 0$, so daß

$$c(s_i, s_j) := -\gamma(s_i, s_j) / \gamma^*(s_j) \geq c' \quad \text{für alle } s_i \in S1_j$$

Wegen [2.5] existiert ein $t_0 \in T$, so daß für alle $t \geq t_0$:

$$\alpha_t(S2_j | s_j, y_k) \leq 1/2 \cdot c' \cdot \alpha_t(S1_j | s_j, y_k) \quad \text{für alle } y_k \in Y$$

d.h. aber, daß

$$\begin{aligned} [2.6] \quad \Sigma(\hat{c}_t | s_j, y_k) &= \sum_{S1_j} c(s_i, s_j) \cdot \alpha_t(s_i | s_j, y_k) + \\ &\quad + \sum_{S2_j} c(s_i, s_j) \cdot \alpha_t(s_i | s_j, y_k) \\ &\geq c' \cdot \alpha_t(S1_j | s_j, y_k) - \alpha_t(S2_j | s_j, y_k) \\ &\geq c'/2 \cdot \alpha_t(S1_j | s_j, y_k) \\ &\geq c'/4 =: c_1 > 0 \quad \text{für } t \geq t_0, y_k \in Y1 \end{aligned}$$

Im folgenden nehmen wir an, daß $Y_2 := \overset{\circ}{Y} \setminus Y_1 \neq \emptyset$ ist, weil sonst die Behauptung ohnehin trivialerweise erfüllt ist.

Beh.(1): Für den zu (s_j, Y_1) gehörigen Modifikationsprozeß $\hat{\mathcal{P}}_t := \hat{\mathcal{P}}_t(j)$ gilt: $\{\hat{\mathcal{P}}_t, t \in T\}$ ist ein Semimartingal, d.h.

$$\mathbb{E}(\hat{\mathcal{P}}_{t+1} \mid \hat{\mathbb{E}}_\tau (\tau \leq t), \hat{S}_t) \geq \hat{\mathcal{P}}_t \quad \pi\text{-fast sicher}$$

Bew.: Nach (1.22) gilt:

$$\begin{aligned} \pi(\{\hat{\mathcal{P}}_t > \mathbb{E}(\hat{\mathcal{P}}_{t+1} \mid \hat{\mathbb{E}}_\tau (\tau \leq t), \hat{S}_t)\}) &= \\ &= \pi(\{\mathbb{E}(\hat{\mathcal{P}}_{t+1} \mid \hat{\mathbb{E}}_\tau (\tau \leq t), \hat{S}_t) < 0\}). \end{aligned}$$

Weiter folgt nach (1.22d) für $\hat{Y}_t = y_{kt}$, $h=1,2$:

$$\begin{aligned} \mathbb{E}(\hat{\mathcal{P}}_t \mid \hat{\mathbb{E}}_\tau (\tau \leq t), \hat{S}_t) &= \\ &= \sum_{y_k \in Y_h} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in \hat{Y}} \min_{t,kt} \cdot \hat{p}_{t,kt} \cdot \hat{S}_t(k,kt) \cdot \mathbb{E}(\hat{c}_t \mid s_j, y_{kt}) \\ &\quad + 0 \end{aligned}$$

mit $\hat{p}_{t,k} := \hat{\mu}_t(y_k \mid s_j)$, $\min_{t,kt} := 2 \cdot \min(\hat{p}_{t,k}; 1 - \hat{p}_{t,k})$ und $\hat{r}_{t,k} := \hat{p}_{t,k} \cdot (1 - \hat{p}_{t,k})$, wobei $\hat{S}_t(k,kt)$ durch [1.22] festgelegt ist.

Da für $y_{kt} \in Y_1$ und $y_k \in Y_2$ $\hat{S}_t(k,kt) < 0$, ferner für $y_{kt} \in Y_1$ $\mathbb{E}(\hat{c}_t \mid s_j, y_{kt}) > 0$ und nach Definition von Y_1, Y_2 demzufolge

$$\sum_{y_k \in Y_2} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \min_{t,kt} \cdot \hat{S}_t(k,kt) \cdot \mathbb{E}(\hat{c}_t \mid s_j, y_{kt}) \cdot \hat{p}_{t,kt} < 0$$

ergibt sich wegen (1.22 f):

$$\sum_{y_k \in Y_1} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \hat{g}_t(k, kt) \cdot \xi(\hat{c}_t | s_j, y_{kt}) \cdot \hat{p}_{t,kt} \cdot \hat{m}_{t,kt}^{in} > 0$$

$$\begin{aligned} \mathbb{E}(\{\hat{\eta}_t | \hat{E}_\tau (\tau < t), \hat{S}_t = s_j\}) &= \\ &= \sum_{y_k \in Y_1} \hat{r}_{t,k} \cdot \left\{ \sum_{y_{kt} \in Y_1} \hat{g}_t(k, kt) \cdot \xi(\hat{c}_t | s_j, y_{kt}) \cdot \hat{p}_{t,kt} \cdot \hat{m}_{t,kt}^{in} \right. \\ &\quad \left. + \sum_{y_{kt} \in Y_2} \hat{g}_t(k, kt) \cdot \xi(\hat{c}_t | s_j, y_{kt}) \cdot \hat{p}_{t,kt} \cdot \hat{m}_{t,kt}^{in} \right\} \end{aligned}$$

Für $y_k \in Y_1$ und $y_{kt} \in Y_2$ gilt: $\hat{g}_t(k, kt) < 0$ und für $y_{kt} \in Y_2$ ist $\xi(\hat{c}_t | s_j, y_{kt}) \leq 0$. Darum ist folgende Abschätzung möglich:

$$\mathbb{E}(\{\hat{\eta}_t | \hat{E}_\tau (\tau < t), \hat{S}_t = s_j\}) > \{0 + 0\} = 0$$

π -fast sicher, d.h. $\{\hat{\mathbb{P}}_t, t > 0\}$ ist ein Semimartingal.

Da weiterhin $\sup \mathbb{E}|\hat{\mathbb{P}}_t| < \infty$, folgt nach DOOB(1965;324):

Es existiert ein $\hat{\mathbb{P}}_\infty$ mit $\pi\{\lim_{t \rightarrow \infty} \hat{\mathbb{P}}_t = \hat{\mathbb{P}}_\infty\} = 1$.

Annahme (a): Sei $\hat{\mathbb{P}}_\infty = \begin{cases} 0 & \text{mit Wahrsch. } \alpha > 0 \\ 1 & \text{mit Wahrsch. } 1-\alpha > 0 \end{cases}$

Dann gibt es nach Voraussetzung [2.5] zwei reihenkonvergente Nullfolgen $(\varepsilon'_t)_{t>0}$, $(\varepsilon''_t)_{t>0}$, so daß gemäß dem Lernalgorithmus für $\hat{\mathbb{P}}_{t-1} = a \in (0,1)$ gilt:

$$\begin{aligned} \hat{\mathbb{P}}_t &\geq a + 2c' \cdot r_{t, \min(a, 1-a)} \text{ mit Wahrsch. } a \cdot x_t(S1_j | s_j, Y1) \geq a \cdot (1 - \varepsilon'_t) \\ \hat{\mathbb{P}}_t &\geq a & " & (1-a) \cdot x_t(S2_j | s_j, Y2) \geq (1-a) \cdot (1 - \varepsilon''_t) \\ \hat{\mathbb{P}}_t &\leq a - 2c' \cdot r_{t, \min(a, 1-a)} & " & (1-a) \cdot x_t(S1_j | s_j, Y2) \leq (1-a) \cdot \varepsilon''_t \\ \hat{\mathbb{P}}_t &\leq a & " & a \cdot x_t(S2_j | s_j, Y1) \leq a \cdot \varepsilon'_t \end{aligned}$$

d.h. es gibt eine Nullfolge $(\varepsilon_t)_{t>0}$ mit $\sum_{t=1}^{\infty} \varepsilon_t < \infty$,
so daß

$$[2.7] \quad \hat{\mathcal{P}}_{1_t} < \hat{\mathcal{P}}_{1_{t-1}} \quad \text{mit Wahrscheinlichkeit} \leq \varepsilon_t.$$

Nach dem Satz von Borel-Cantelli (vgl. RICHTER, 1966; 400)
treten dann die Alternativen [2.7] π -fast sicher nur
endlich oft auf, d.h.

$\hat{\mathcal{P}}_{1_{\infty}} > 0$ π -fast sicher. Dies ist ein Widerspruch
zur Annahme (a).

Annahme (b): $\lim_{t \rightarrow \infty} \sum \hat{\mathcal{P}}_{1_t} = \alpha$ mit $0 < \alpha < 1$.
 $\circ \text{EdA } \alpha \leq 1/2$

Wegen $0 < \hat{\mathcal{P}}_{1_t} < 1$ für $t \in T$ folgt dann für
 $\alpha'_t := \sum \hat{\mathcal{P}}_{1_t}^4$ und $\alpha''_t := \sum \hat{\mathcal{P}}_{1_t}^5$: $0 < \alpha''_t < \alpha'_t < 1$ ($t \in T$)
und wegen Ungültigkeit der Annahme (a) folgt auch:

$$\beta := \lim_{t \rightarrow \infty} (\alpha'_t - \alpha''_t) > 0.$$

Für $n_1 :=$ Zahl der y_k in Y_1 gibt es eine natürliche Zahl k
derart groß, daß die Ungleichung

$$k \geq 2 \cdot n_1 / (c' \cdot \beta) \quad \text{erfüllt ist, wobei } c' \text{ durch [2.6]}$$

gegeben ist. Sei nun $\varepsilon := 1/k$ gegeben, dann existiert ein
 $t_0 \in T$, so daß für alle $t \geq t_0$ gilt:

$$[2.8] \quad \begin{cases} (\alpha'_t - \alpha''_t) > \beta/2 & \text{und} \\ \alpha - \varepsilon \leq \sum \hat{\mathcal{P}}_{1_t} \leq \alpha & t \in T, \text{ wegen der Monotonie von } \hat{\mathcal{P}}_{1_t} \end{cases}$$

Andererseits ist aber für $t \geq t_0$: $t \in T^* = \{t \in T \mid \hat{S}_t = s_j\}$

$$\mathcal{E}(\{ \sum(\{ \hat{\mathcal{P}}_{1_t} \mid \hat{E}_t(z < t), \hat{S}_t \}) \}) =$$

$$\begin{aligned}
 &\geq \sum_{y_k \in Y_1} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \hat{p}_{t,kt}^2 \cdot \hat{f}_t(k, kt) \cdot \sum (\hat{c}_t | s_j, y_{kt})) \\
 &= \sum_{y_k \in Y_2} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \hat{p}_{t,kt}^2 \cdot \underbrace{(-\hat{f}_t(k, kt))}_{> 0} \cdot \underbrace{\sum (\hat{c}_t | s_j, y_{kt}))}_{> 0} \\
 &> \sum (c' \cdot \sum_{y_k \in Y_2} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} (-\hat{f}_t(k, kt)) \cdot \hat{p}_{t,kt}^2) \\
 &= \sum (c' \cdot \sum_{y_k \in Y_2} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \frac{\hat{p}_{t,kt}^3 \cdot (1 - \hat{p}_{t,kt})}{\sum_{x+kt} \hat{r}_{t,x}}) \\
 &\geq \sum (c' \cdot \sum_{y_k \in Y_2} \hat{r}_{t,k} \cdot \sum_{y_{kt} \in Y_1} \hat{p}_{t,kt}^3) \\
 &\geq \sum (\frac{c'}{n_1} \cdot \hat{\mathcal{P}}_t^3 \cdot \sum_{y_k \in Y_2} \hat{r}_{t,k}), \quad \text{weil } \sum_{y_{kt} \in Y_1} p_{t,kt}^3 \geq \frac{\hat{\mathcal{P}}_t^3}{n_1} \\
 &\geq \sum (\frac{c'}{n_1} \cdot \hat{\mathcal{P}}_t^4 \cdot (1 - \hat{\mathcal{P}}_t)) \\
 &\geq c' / n_1 \cdot (\alpha'_t - \alpha''_t) \\
 &> c' \cdot \beta / (2 \cdot n_1) \geq \varepsilon,
 \end{aligned}$$

d.h. $\sum (\hat{\mathcal{P}}_{t+1}) = \sum (\{ \sum (\{ \hat{\mathcal{P}}_{t+1} | \hat{E}_t(\tau \leq t), \hat{S}_t \cdot \}) \}) > \alpha$

Das ist aber ein Widerspruch zu [2.8]. Mithin folgt:

$\lim_{t \rightarrow \infty} \sum \hat{\mathcal{P}}_t = 1$. Nach (1.26) können wir schließen:

$\pi \lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t = 1$. Wegen der starken Konvergenz von $\hat{\mathcal{P}}_t$ und wegen (1.26d) impliziert dies: $\pi (\lim_{t \rightarrow \infty} \hat{\mathcal{P}}_t = 1) = 1$.

Da Y_1 die Bedingung [2.2b] erfüllt, folgt die Behauptung. ./.

(2.5) Korollar:

$\tilde{\gamma}$ sei das Raster eines unabhängigen LS^H und $\tilde{\gamma} \in \tilde{\gamma}^*$ sei ein wesentlicher Rasterausschnitt mit $\tilde{\gamma}^* \neq \emptyset$ sowie $\tilde{\gamma}^*$ quasideterministisch für alle $t \geq t_0$ ($t_0 \in T$). Für jedes $s_j \in \tilde{\gamma}^*$ konvergiert $Y1_t(j) := \{y_k \in \tilde{\gamma} \mid \sum(\hat{c}_t | s_j, y_k) > 0\}$ mit $\lim_{t \rightarrow \infty} Y1_t(j) \neq \emptyset$.
 Beh.: A approximiert auf $\tilde{\gamma}$ ein HS.

Beweis: Sei $s_j \in \tilde{\gamma}^*$. Dann gibt es ein $y_k \in \tilde{\gamma}$ und ein $t_1 \geq t_0$, so daß $\sum(\hat{c}_{t_1} | s_j, y_k) > 0$ und wegen der Quasideterminiertheit von

$\tilde{\gamma}^*$ ist dann auch $\hat{p}_t(s_j, y_k) < 0$ $\forall$ -fast sicher für alle $t \geq t_1$, d.h. aber für $Y1_{t_0} := \{y_k \in \tilde{\gamma} \mid \sum(\hat{c}_t | s_j, y_k) > 0\}$ folgt:

$$\alpha_t(S2_j | s_j, Y1_{t_0}) = 0 \text{ für alle } t \geq t_1 \text{ und } S2_j := \{s_i \in \tilde{\gamma} \mid \gamma(s_i, s_j) \geq 0\}.$$

Ebenso folgt aus $\sum(\hat{c}_t | s_j, y_k) \leq 0$, daß $\hat{p}_t(s_j, y_k) \geq 0$ $\forall$ -fast sicher und daher ist

$$\alpha_t(S1_j | s_j, Y2_{t_0}) = 0 \text{ für alle } t \geq t_1 \text{ und } S1_j := \tilde{\gamma} \setminus S2_j \text{ sowie } Y2_{t_0} = \tilde{\gamma} \setminus Y1_{t_0}.$$

Da aber $\lim_{t \rightarrow \infty} Y1_t(j) \neq \emptyset$, folgt nach (2.4) die Behauptung.

(2.6) Korollar:

$\tilde{\gamma}$ sei das Raster eines unabh. LS und $\tilde{\gamma}^* \neq \emptyset$ sei quasistabiler, wesentlicher Rasterausschnitt. Für jedes $s_j \in \tilde{\gamma}^*$ existiert $Y1(j) := \{y_k \in \tilde{\gamma} \mid \sum(\hat{c} | s_j, y_k) > 0\} \neq \emptyset$ mit der Eigenschaft, daß

$$[2.8] \quad \sum_{y \in Y1(j)} \alpha(S2_j | s_j, y_k) + \sum_{y \in Y2(j)} \alpha(S1_j | s_j, y_k) = 0$$

wobei $Y2(j) := \tilde{\gamma} \setminus Y1(j)$, $S1_j, S2_j$ wie in (2.4) definiert.

Beh.: A approximiert auf $\tilde{\gamma}$ ein HS.

Beweis: Aus [2.8] folgt sofort [2.5]. Daraus ergibt sich nach (2.4) die Behauptung. ./.

Bemerkung: Satz (2.3) folgt nun als Spezialfall von (2.5) oder auch als Spezialfall von (2.6).

(2.7) Korollar:

A sei ein unabhängiges LS und $\tilde{\mathcal{V}}$ sei ein wesentliches Situationsraster, wobei alle $s_j \in \tilde{\mathcal{V}}$ die Bedingungen erfüllen:

[2.9a] $\gamma(s_i, s_j) \leq 0$ für alle $s_i \in \tilde{\mathcal{V}}$

[2.9b] $Y1_t := \{y_k \in \tilde{\mathcal{Y}} \mid \sum \hat{c}_t | s_j, y_k \rangle > 0\}$ wird schließlich konstant $= Y1 \neq \emptyset$ und

[2.9c] $\sum_{t>0} \alpha_t(S2_j | s_j, Y1_t) < \infty$

Beh.: A approximiert auf $\tilde{\mathcal{V}}$ ein HS.

Beweis: Da $\sum \hat{c}_t | s_j, y_k \rangle \leq 0$ genau dann, wenn

$\alpha_t(S1_j | s_j, y_k) = 0$, folgt für $Y2_t := \tilde{\mathcal{Y}} \setminus Y1_t$:

$Y2_t = \{y_k \in \tilde{\mathcal{Y}} \mid \alpha_t(S1_j | s_j, y_k) = 0\}$ für alle $s_j \in \tilde{\mathcal{V}}$. d.h. aber, die Voraussetzungen von (2.4) sind erfüllt. ./.

2.2 Approximationsbedingungen für abhängige LS

Ähnlich wie im Fall unabhängiger LS lassen sich auch für abhängige LS Aussagen über die Approximation an Homöostatische Systeme machen. Wir wollen in diesem Abschnitt jedoch keine mathematisch tieferliegende Sätze aufstellen, sondern lediglich einen für die Anwendungen im psychologischen Bereich interessanten Fall herausgreifen, der die Eignung des Modells für verschiedenartige praktische Problemstellungen verdeutlichen kann.

Im folgenden werden die Lernprozesse eines abhängigen LS untersucht, wenn die Umwelt in einer ganz bestimmten Weise manipulativ auf die Reaktionen des LS eingeht. Inhaltlich gesehen, lassen sich solche Umweltreaktionen $(f_t, t \in T)$ als Verhaltensformungsexperimente bezeichnen.

(2.8) Def.: Verhaltensformungsexperiment (Vf.exp.)

Sei A ein abh. LS mit $\mathcal{Y} = \{y_k = (a_{k1}, \dots, a_{kn}) \mid a_k = 0, 1\}$, n, n_1, r natürliche Zahlen mit $n_1 \leq n$ und $s_j \in \mathcal{S}$ mit $\|s_j\| > 0$, sowie $Y_1(1) := \{y_k \in \mathcal{Y} \mid a_{kv} = 1 \ (v=1, \dots, 1) \text{ und } Y_2(1) := \mathcal{Y} \setminus Y_1(1)$. $S_{1j} := \{s_i \in \mathcal{S} \mid \gamma(s_i, s_j) < 0 \text{ und } S_{0j} := \{s_i \in \mathcal{S} \mid \gamma(s_i, s_j) = 0\}$.

Dann heißt $(f_t, t \in T)$ Verhaltensformungsexp. r -ter Ordnung für $(s_j; \hat{a}_1, \dots, \hat{a}_{n_1})$, wenn

$$(1) \quad f_0(s_j) = [1]$$

$$(2) \quad \text{Für } t \in T^* := \{t > 0, \hat{S}_t = s_j\} \text{ mit } f_t(s_j) = [1] \text{ gilt:}$$

$$[2.10a] \quad f_{t+1}(s_j) = \begin{cases} [1+1] & , \text{ falls } Y_{t+1} \in Y_1(1) \ (\gamma = 0, \dots, r), 1 < n_1 \\ [1] & \text{sonst} \end{cases}$$

$$\text{Dabei ist } f_t(s_j) = [1] : \begin{cases} x_{t+1}(S_{1j} | s_j, Y_1(1)) = 1 \text{ und} \\ x_{t+1}(S_{0j} | s_j, Y_2(1)) = 1 \text{ ist.} \end{cases}$$

(2.9) Satz: Approximationssatz für abhängige LS

Sei A ein abhängiges LS über $\mathcal{F}$, $\tilde{\mathcal{T}}$ sei wesentliches, unabhängiges Situationsraster. Für alle $s_j \in \tilde{\mathcal{T}}$ existieren natürliche Zahlen $n_{1,r}$, so daß $(f_t, t > 0)$ ein Verhaltensformungsexperiment r -ter Ordnung für $(s_j; \hat{a}_1, \dots, \hat{a}_r)$ ist.

Beh.: a) Es gibt $\mathcal{W}$ -fast sicher ein $t_0 = t_0(s_j)$, so daß für alle $t \geq t_0$ und für alle $y_k \in Y_1(n_1)$ gilt:

$$x_{t+1}(S_{1j} | s_j, y_k) = 1$$

b) Für den zu $(s_j, Y_1(n_1))$ gehörenden Modifikationsprozeß

$\{\hat{S}^1_t, t \geq 0\}$ gilt: $\mathbb{P}(\lim_{t \rightarrow \infty} \hat{S}^1_t = 1) = 1$, d.h.

A approximiert auf $\tilde{\mathcal{T}}$ ein HS.

Beweis: a) Da wegen [2.10a] $f_t(s_j) = [1_t]$ mit $(1_t)_{t=1,2,\dots}$ eine nicht fallende Folge ist, genügt es wegen der Definition eines Verhaltensformungsexp. zu zeigen: Es gibt $\mathcal{W}$ -fast sicher ein $t_0 \in T$, so daß $1_{t_0} = n_1$ ist.

Annahme 1: $1_t = 1$ für alle $t \in T$. Dann gilt auch für $t \in T^*$ und

für $\hat{r}_{t,1} := \hat{p}_{t,1} \cdot (1 - \hat{p}_{t,1})$ wegen [1.25]:

$$\hat{\mathcal{P}}^1(1)_t = \begin{cases} \hat{c}_t \cdot \hat{m}_{t,1} \cdot \hat{r}_{t,1} \geq c' \cdot r_{t,1}^2 > 0 & \text{für } \hat{Y}_t \in Y_1(1) \\ = 0 & \text{sonst} \end{cases}$$

d.h. $\hat{\mathcal{P}}^1(1)_t \geq \hat{\mathcal{P}}^1(1)_0$ $\mathcal{W}$ -fast sicher. d.h. für jedes r , nat. gibt es nach DOOB (1965; 324) $\mathcal{W}$ -fast sicher ein $t_1 \in T^*$, so daß $\hat{Y}_{t_1 - \xi} \in Y_1(1)$ ($\xi = 0, \dots, r$) und daher $f_{t_1}(s_j) = [2]$. Das ist ein Widerspruch zur Annahme 1.

Annahme 2: Es gibt ein $1 < l < n_1$, so daß $\mathcal{W}$ -fast $f_t = [1]$ für alle $t \geq t_1$.

Nun ist $\hat{y}^1(1)_t = \prod_{v=1}^t \hat{p}_{t+1,v} - \prod_{v=1}^t \hat{p}_{t,v}$ nach [1.24] und [1.25]

für $t \geq t_1$:

$$[2.11] \quad \hat{y}^1(1)_t = \begin{cases} \geq c' \cdot \sum_{v=1}^t \hat{r}_{t,v}^2 \prod_{1 \leq v \leq t} \hat{p}_{t,v} > 0 & \text{mit W. } \prod_{v=1}^t \hat{p}_{t,v} > 0 \\ \geq 0 & \text{sonst} \end{cases}$$

mit $\prod_{v=1}^t \hat{p}_{t,v} = \mu_t(\hat{a}_1 = 1 | s_j) \cdot \mu_t(\hat{a}_2 = 1 | s_j, \hat{a}_1 = 1) \dots \mu_t(\hat{a}_r = 1 | \cdot) \text{ kgf gegen } 1$

Wie oben folgt dann nach DOOB (1965;324), daß $Y_{t_i-j} \in Y(1)$ ($j = 0, \dots, r$)

für $t_i \in T$ $\bar{W}$ -fast sicher, so daß die Behauptung a) folgt.

b) Wie bei [2.11] folgt, daß $\hat{y}^1(n1)_t > 0$ mit einer positiven, nicht fallenden Wahrscheinlichkeit und sonst = 0 ist, sofern $f_t(s_j) = [n_1]$ schon erreicht ist, so daß nach DOOB (1965;324) $\hat{y}$ -fast sicher

$\hat{y}^1(n1)_t$ unendlich oft erhöht wird, d.h. $\lim_{t \rightarrow \infty} \hat{y}^1(n1)_t = 1$ ist

und somit die Behauptung bewiesen ist. ./.

Bemerkung: Wie man an dem Lernalgorithmus mit Hilfe von Beispielen sofort feststellt, ist die Konvergenzgeschwindigkeitserwartung von der Reihenfolge der Komponenten (wenn diese verschiedene Ausgangswahrscheinlichkeiten besitzen) und von der Wahl von r abhängig. Je höher r gewählt wird, um so konservativer ist die Einstellung des Experimentators zum LS. Bei extrem kleinem r besteht die Gefahr, daß die Stufe $f_t(s_j) = n_1$ bereits erreicht wird, wenn noch die bed. Verhaltenswahrscheinlichkeiten der einzelnen Komponenten und damit erst recht die Produktwahrscheinlichkeiten gering sind. Bei sehr hohem r findet in der Regel ein Überlernen statt, was mit einem Motivationsverlust z.B. verbunden sein dürfte.

(2.10) Korollar:

Sei A ein abhängiges LS über $\tilde{\mathcal{E}}$ und $\tilde{y} = (a_{k1}, \dots, a_{kn}) : a_{kv} = 0, 1$

$\tilde{\mathcal{E}}$ sei ein Binärgitter und es existiert ein $s_j \in \tilde{\mathcal{E}}$, das die

Voraussetzungen von (2.9) erfüllt. Für alle $s_i \in \tilde{\mathcal{E}} :=$

$\{s_i \in \tilde{\mathcal{E}} \mid \sigma(s_i, s_j) > 0\}$ existiert ein $l = l(s)$ mit

$1 \leq l \leq n_1$, so daß $(f_t, t \geq 0)$ Verhaltensformungsexperiment m -ter Ordnung für $(s_i; \hat{a}_1, \dots, \hat{a}_1)$ ist.

Beh.: Für die zu $(s_i, \hat{a}_1(t))$ gehörigen Modifikationsprozesse $\hat{S}_t(i)$ gilt: $\pi(\lim_{t \rightarrow \infty} \hat{S}_t(i) = 1) = 1$ für alle $s_i \in \tilde{S}$

Beweis: Da gemäß [1.20] und Voraussetzung (2.10) verwandte Situationen von $s_i \in \tilde{S}^*$ einzelne Komponenten allenfalls in gleicher Richtung, lediglich mit einem Faktor kleiner als 1 verstärken, wird durch den Transfer (über das Binärgitter) der Lernprozeß, der nach (2.9) für die s_i im unabhängigen Fall in der gewünschten Richtung konvergiert, durch die Struktur des Binärgitters allenfalls beschleunigt, d.h. $\hat{S}_t(i)$ konvergiert für jedes der s_i stark nach 1. Daraus folgt die Behauptung. ./.

Bemerkung: Auch hier ist die Zunahme der Konvergenzgeschwindigkeit interessant, mehr noch aber die Tatsache, daß nicht mehr die Wesentlichkeit von $\tilde{S}$ vorausgesetzt werden muß. 'Wenn wenigstens für s_j (wie dies vorausgesetzt ist) die Wesentlichkeit gilt, werden die entsprechenden Komponenten auch bei den verwandten Situationen verstärkt, wenn auch um so langsamer, je geringer die Verwandtschaft ist.

Verallgemeinerungen dieser Aussagen sind möglich, die auch mathematisch von Interesse wären. Um jedoch die Brauchbarkeit des Modells für die verschiedenen psychologischen Fragestellungen zu demonstrieren, wenden wir uns nun Verhaltenssystemen mit nicht-konstanten Ereignisrastern zu.

2.3 Strukturapproximation in differentiellen Lernautomaten

Im folgenden wird nicht mehr angenommen, daß das Ereignisraster $\{t\}$ über alle $t \in T$ konstant sei. Die Zahl sinnvoller Festlegungen für f_1 und f_2 kann hier nicht annähernd behandelt werden. Wir reduzieren daher die Untersuchung auf einstufige Lernsysteme und greifen ein für die Anwendungen nützlich Beispiel heraus, das eine typische Möglichkeit des differentiellen Automaten aufzeigt.

Aus Gründen des Aufwands verzichten wir auch auf die Betrachtung von Systemen, die das Situationsraster verfeinern oder gar eine Differenzierung von Verhaltens- und Situationsraster vornehmen. Sie können mit einem ähnlichen Ansatz wie dem angegebenen gelöst werden, sprengen aber wegen der dabei erforderlichen Komplexität den Rahmen der vorliegenden Arbeit. Ebenso lassen wir auch die Betrachtung von Rastervergrößerungen fort, die lediglich der Ökonomie der Informationsspeicherung und der Erhöhung der Konvergenzgeschwindigkeit dienen. (vgl. dazu § 3.2). Dieser und der nächste Abschnitt haben vor allem die Aufgabe aufzuzeigen, welche Möglichkeiten das von uns konzipierte Lernmodell besitzt.

Ziel der Verhaltensdifferenzierung ist die Verfeinerung eines gegebenen Verhaltensrasters $\{y_t\}$ ($t \in T$), bis mindestens eine der verfügbaren Alternativen eine befriedigende Ungleichgewichtsreduktion ermöglicht. Dazu definieren wir zunächst die Abbildung f_1 , die die Bedingungen und den Ablauf der Rasterdifferenzierung festlegt.

(2.11) Def.: Verhaltensdifferentielles LS

Sei $A = (A_t(v), t \in T)$ ein LS über $\dot{f}_t = \dot{f}^* \dot{y}_t$ und $0 \leq \varepsilon_A < 1$, $2 \leq n_A \leq v$. A heißt VERHALTENSDIFFERENTIELLES LS mit der verh.diff. Abb. $f_2 = f_2(n_A, \varepsilon_A)$, wenn gilt: Für alle $\hat{E}_t = e_{it}$ ex. ein $T^{**} \subset T$, so daß

$$\dot{y}_{t+1} = f_2(\dot{y}_t) = \begin{cases} \{y_{1t}, \dots, y'_{kt}, y''_{kt}, \dots, y_{rt}\} & \text{für } t \in T^{**} \\ \dot{y}_t & \text{sonst} \end{cases}$$

mit $y_k = y'_{kt} + y''_{kt}$, wobei der Schnitt allein durch y_k eindeutig bestimmt ist. Dabei ist $t \in T^{**}$, wenn

$$\hat{E}_t = (s_j, y_k) \text{ mit}$$

$$[2.13] \begin{cases} a) \ n(t) := \max (\bar{r} \mid \dot{y}_{t-\bar{r}} = \dot{y}_{t-\bar{r},4}, 0 < \bar{r} \leq v) \geq n_A \\ b) \mid \bar{\xi}(\hat{c}_t \mid s_j, y_k) - c_t(\hat{s}_{t+1}, \hat{s}_t) \mid > \varepsilon_A \end{cases}$$

wobei $\bar{\xi}(\hat{c}_0 \mid s_j, y_k) = c_{jk} \in [-1, +1]$ für $s_j \in \hat{s}^0$ und $y_k \in \dot{y}_v$ sowie

$$[2.14] \quad \bar{\xi}(\hat{c}_{t+1} \mid s_j, y_k) := \begin{cases} 1/n_A [(n_A - 1) \cdot \bar{\xi}(\hat{c}_t \mid s_j, y_k) + c_t(\hat{s}_{t+1}, \hat{s}_t)] & \text{für } \hat{E}_t = (s_j, y_k) \\ \bar{\xi}(\hat{c}_t \mid s_j, \tilde{y}_k) & \text{sonst, für } y_k < \tilde{y}_k \in \dot{y}_t \end{cases}$$

Die Menge der durch endliche Teilung aus $\dot{y}_0$ entstehenden atomaren Zerlegungen von Y bezeichnen wir mit

$$[2.15] \quad \dot{Y}(\dot{y}_0, q_2) := \left\{ \dot{y}_w \in \dot{Y}(Y) : \text{es existiert eine natürliche Zahl } w, \text{ so daß } \dot{y}_w \text{ durch Teilung eines Elements aus } \dot{y}_{\tau-1} \text{ entsteht} \right. \\ \left. (\tau = 1, \dots, w) \right\}$$

Bemerkung: Natürlich gilt wegen (1.2)₍₂₎ und (1.3)_(4a), daß für $y_k = y'_k + y''_k$ mit $y'_k, y''_k \in \check{Y}_{t+1}$

[2.12] $\mu_{t+1}(y'_k | s_j) > 0$ und $\mu_{t+1}(y''_k | s_j) > 0$
sein muß.

(2.12) Satz: Approximationssatz für verh.diff. LS

Sei A ein unabhängiges, verh.diff. LS mit dem Ereignisraster $\check{f}_t = \check{f}^* \check{f}_t$ ($t \in T$) und der verh.diff.Abb. $f_2(n_A, \epsilon_A)$ über $\tilde{\sigma} \subset \check{\sigma}$, wobei gilt:

(a) $\tilde{\sigma}$ ist wesentliches Situationsraster mit

[2.16] $0 < \epsilon_A < 1/2 \cdot \min_{s_i \in \tilde{\sigma}} |c(s_i, s_j)|$ für $s_j \in \tilde{\sigma}$
 $|s_i| \neq |s_j|$

(b) Für alle $s_j \in \tilde{\sigma}$ ex. ein $\check{y}^* \in \check{Y}(\check{y}_0, f_2)$, so daß für alle $y_k \in \check{y}^*$ gilt: Es gibt ein $s_k = s_i(y_k) \in \check{\sigma}$:

[2.17] $\overline{\sum_{t=1}} \sup_{y \in y_k} \overline{\sum_{s_i \in \tilde{\sigma}}} \alpha_t(s_i | s_j, y) < \infty$

(c) Für alle $s_j \in \tilde{\sigma}$ ex. ein $t_0 \in T$:

[2.18] $Y1_t(s_j, \check{y}^*) := \{y_k \in \check{y}^* | \mathbb{E}(\hat{c}_t | s_j, y_k) > 0\} = Y1 \neq \emptyset$
für $t \geq t_0$

Beh.: Für jedes $s_j \in \tilde{\sigma}$ gilt:

a) Es gibt π -fast sicher ein $t_1 \in T$, so daß für alle $t \geq t_1$ $\check{y}_{t,1}^0 := \{y_k \in \check{y}^* | \overline{\sum_{t=1}} \alpha_t(s_2 | s_j, y_k) < \infty\} = \check{y}_{t_1,1}^0$

b) Der zu $(s_j, \check{y}_{t_1,1}^0)$ gehörende Modifikationsprozeß $\{\hat{P}_t(j), t > t_1\}$ konvergiert π -fast sicher nach 1,

d.h. A approximiert auf $\tilde{\sigma}$ ein HS.

Beweis: Sei $s_j \in \tilde{\mathcal{S}}$ und eine Folge $(\hat{E}_t = e_{it}, t > 0)$ gegeben.

Für die auf diese Weise eindeutigen Realisierungen $\tilde{y}_t (t \geq 0)$ betrachten wir die disjunkte Zerlegung

$$\tilde{y}_t = \tilde{y}_{t,0} + \tilde{y}_{t,1} + \tilde{y}_{t,2} \quad \text{mit}$$

$$\tilde{y}_{t,1} := \{ y_k \in \tilde{y}_t \mid \sum_{\tau \geq t} \alpha_\tau(S2_j | s_j, y_k) < \infty \} \quad \text{und}$$

$$\tilde{y}_{t,2} := \{ y_k \in \tilde{y}_t \mid \sum_{\tau \geq t} \alpha_\tau(S1_j | s_j, y_k) < \infty \}$$

wobei $S1_j := \{ s_i \in \tilde{\mathcal{S}} \mid c(s_i, s_j) > 0 \}$ und $S2_j = \tilde{\mathcal{S}} \setminus S1_j$.

Natürlich folgt, daß

$$\tilde{y}_{t,0} = \{ y_k \in \tilde{y}_t \mid \sum_{\tau \geq t} \alpha_\tau(S2_j | s_j, y_k) = \sum_{\tau \geq t} \alpha_\tau(S1_j | s_j, y_k) = \infty \}$$

Annahme(A): $\left| \begin{array}{l} \tilde{y}_{t,0} \neq \emptyset \text{ für alle } t \in T \text{ und es gibt ein } t_0 \in T \text{ mit} \\ \tilde{y}_{t,0} = \tilde{y}_{t_0,0} \text{ für alle } t \geq t_0. \end{array} \right.$

Sei $T^* := (t_i)_{i=1,2,\dots} = \{ t \in T \mid \hat{S}_t = s_j \}$ und $y_k \in \tilde{y}_{t_0,0}$.

Dann existiert nach Vor. (b) eine Zerlegung $y_k = \sum_{\lambda=1}^p y_k^{(\lambda)}$

mit $1 \geq 2$ und oEdA

$$\sum_{t > 0} \alpha_t(S2_j | s_j, y_k^{(1)}) < \infty \quad \text{sowie}$$

$$\sum_{t > 0} \alpha_t(S1_j | s_j, y_k^{(2)}) < \infty \quad p\text{-fast sicher.}$$

Nach dem Lemma (1) von BOREL-CANTELLI (vgl. RICHTER, 1966; 400)

existiert dann p -fast sicher ein $n_1 \in T$: so daß für alle $t > n_1$:

$\hat{Y}_t \in y_k^{(2)}$ impliziert $\hat{S}_{t+1} \in S2_j$ und $\hat{Y}_t \in y_k^{(1)}$ impliziert $\hat{S}_{t+1} \in S1_j$,

so daß gemäß dem Lernalgorithmus in diesem Fall wegen der λ -Stetigkeit des Maßes μ gilt:

$$\mu'_{t+1}(y_k^{(i)} | s_j) > \mu_t(y_k | s_j) \text{ für } \hat{E}_t = (s_j, y_k) \text{ (s. (1.3))}.$$

Sei nun $r > 0$, natürlich und

$$[2.19a] \quad M_r := \{ \hat{S}_{t_r} = s_j, \hat{Y}_{t_r} = y_k^{(i)} \}$$

Dann gilt für die bedingte Wahrscheinlichkeit

$$\begin{aligned} [2.19b] \quad \mu_r(M_r | I_{M_1}, \dots, I_{M_{r-1}}) &= \\ &= \mu_{t_r}(\{ \hat{Y}_{t_i} = y_k^{(i)} | E_{t_i} = (s_j, y_k^{(i)}), (i=1, \dots, r-1), \\ &\quad \hat{S}_{t_i} = s_j \}) \cdot \mu_{t_r}(\{ \hat{S}_{t_r} = s_j | \hat{E}_{t_i} = (s_j, y_k^{(i)}) (i=1, \dots, r-1) \}) \\ &\geq \mu_{n_1}(\hat{Y}_t = y_k^{(i)} | s_j) \cdot 1 \quad \text{wegen (2.1)} \\ &> 0 \quad \mu\text{-fast sicher, wegen [2.12], so daß} \end{aligned}$$

$$[2.19c] \quad \sum_{r=1}^{\infty} \mu_r(M_r | I_{M_1}, \dots, I_{M_{r-1}}) = \infty$$

weil für $\hat{S}_t \neq s_j$ $\mu_t(y_k | s_j)$ nicht verändert wird. Nach DOOB (1965;324) tritt dann wegen $s_j \in \tilde{\sigma}$ (wesentlich) $\hat{Y}_t = y_k \supset y_k^{(i)}$ μ -fast sicher unendlich oft ein. Da andererseits unter der Annahme (A) die Wahrscheinlichkeit des Ereignisses

$$[2.20a] \quad M_r^{(h)} := \left\{ \begin{array}{l} \text{beim } r\text{-ten Auftreten von} \\ (s_j, y_k) \text{ tritt } y_k^{(h)} \text{ auf} \end{array} \right\} \quad (h=1, \dots, l)$$

wesentlich von dem Lebesguemaß λ abhängt und nach [2.12] größer als 0 ist, folgt für $h = 1, \dots, l$:

$$[2.20b] \quad \mu_r(M_r^{i_k} | I_{M_1^{i_1}}, \dots, I_{M_{r-1}^{i_{r-1}}}) = \mu_r(M_r^{i_k}) > 0$$

und daher

$$[2.20c] \quad \sum_{r=1}^{\infty} \mu_r(M_r^{i_k} | I_{M_1^{i_1}}, \dots, I_{M_{r-1}^{i_{r-1}}}) = \infty$$

d.h. aber nach dem Vorgenannten und dem Lemma 2 von BOREL-CANTELLI (vgl. RICHTER, 1966; 400) folgt μ -fast sicher: Jedes der $y_k^{i_h}$ ($h=1, \dots, l$) tritt unendlich oft ein.

Nach Vor.[2.16] muß aber dann für hinreichend hohes t [2.13] erfüllt sein, weil entweder

$$|\bar{\varepsilon}(\hat{c}_t | s_j, y_k) - c(\hat{S}_{t+1}^1, s_j)| > \varepsilon_n \quad \text{oder} \\ |\bar{\varepsilon}(\hat{c}_t | s_j, y_k) - c(\hat{S}_{t+1}^2, s_j)| > \varepsilon_n$$

ist, so daß y_k μ -fast sicher in y_k^1 und y_k^2 geteilt wird. Dies ist ein Widerspruch zur Annahme (A).

Wegen der Vor. b), daß nur endlich viele, von der Vergangenheit unabhängige Teilungen erforderlich sind, um

j^* zu erreichen, wird μ -fast sicher schließlich ein n_2 durchlaufen, so daß $j_{t,0}^o = \emptyset$ für alle $t \geq n_2$.

analog zeigt man, daß die Kategorien von $j_{t,1}^o$ so lange geteilt werden, bis schließlich für alle $y_k \in j_{t,1}^o$ gilt:

Es gibt ein $s_k = s_i(y_k)$:

$$\sum_{r>t} \alpha_r(s_i | s_j, y_k) < \infty \quad \text{für alle } s_i \neq s_k$$

weil die gleiche Abschätzung wie in [2.19b] möglich ist.

Bei $\check{y}_{t,2}$ ist eine solche Abschätzung nach unten nicht sicher möglich, aber auch nicht nötig, weil für alle $y_k \in \check{y}_{t,2}$ gilt:

$$[2.21] \quad \hat{\gamma}_{1_t}(j) \geq 0 \quad \pi\text{-fast sicher für } \hat{E}_t = (s_j, y_k),$$

unabhängig davon, ob gegebenenfalls eine unbegrenzte Teilung stattfindet oder nicht.

Für p -fast alle Ereignisfolgen $(e_{it}, t \geq 0)$ kann man also schließen, daß ein (von der Folge abhängiges) t_1 existiert, so daß

$$\check{y}_{t,0} = \emptyset \quad \text{und} \quad \check{y}_{t,1} = \check{y}_{t_1,1} \quad \text{für alle } t \geq t_1;$$

d.h. es gilt die Behauptung a).

Wegen Voraussetzung [2.18] ist aber $\check{y}_{t_1,1} \neq \emptyset$, so daß nach

[2.21] wie im Beweis von (2.3) folgt: Der zu $\check{y}_{t_1,1}$ gehörige Modifikationsprozeß $\hat{\beta}_{1_t}(j)$ konvergiert gegen 1. Da

dieser Schluß aber für p -fast jede Ereignisfolge gilt, folgt die Konvergenzbehauptung π -fast sicher, d.h. β approximiert ein HS; es gilt die Beh.b). ./.

Die Betrachtung differentieller Lernprozesse bei abhängigen LS dürfte im Prinzip analoge Resultate erbringen. Deren Explizierung muß hier jedoch entfallen.

2.4 Modelle mehrstufiger Lernsysteme

Bisher wurde das Zustandsraster $\overset{c}{\mathcal{Z}}$ immer konstant gehalten, d.h. das für die Ungleichgewichtsänderung und somit für die Lernrichtung relevante Raster. Um jedoch Lernprozesse zu erklären, die nicht durch sofortige "Belohnung" (bzw. "Bestrafung") zustandekommen, muß diese Voraussetzung aufgegeben werden. Dies ist in unserem allgemeinen Modell (1.3) eines VS vorgesehen.

Als mehrstufige LS bezeichnen wir LS, die durch Hinzufügen von bedingten Sollwerten, d.h. durch eine Vergrößerung der Dimension des Zustandsraums die Bewertung der Ungleichgewichtsänderung verändern können.

Wegen der Komplexität der hier auftretenden Fragen können wir das Konzept nur grob skizzieren, keineswegs exakt oder gar vollständig behandeln.

In (1.3) wird zugelassen, daß $\overset{c}{\mathcal{Z}}_t$ nicht konstant ist. Für das Konzept eines mehrstufigen LS, d.h. eines LS, das seinen Lernprozeß auf eine nach frühestens 2 Takten auftretende "Belohnung" oder "Bestrafung" ausrichten kann, ist folgender Ansatz zweckmäßig:

Ist $\overset{c}{\mathcal{Z}}_0$ das Zustandsraster über $(Z_0, \overset{c}{\mathcal{Z}}_0)$, dann gilt für jedes $t > 0$ und für jede atomare Zerlegung $\overset{c}{\mathcal{Z}}_t$ über $(Z, \overset{c}{\mathcal{Z}})$ mit $Z := Z_0 * X$: Die Einbettung von $\overset{c}{\mathcal{Z}}_0$ in Z $\overset{c}{\mathcal{Z}}_0^* := Z(\overset{c}{\mathcal{Z}}_0)$ erfüllt die Relation:

$$[2.22] \quad \overset{c}{\mathcal{Z}}_0^* \subset \overset{c}{\mathcal{Z}}_t \quad \text{für alle } t \in T$$

d.h. die Rasterstruktur von $\overset{c}{\mathcal{Z}}_0$ wird nicht angetastet, Es

können höchstens Wahrnehmungsdimensionen in das Raster einbezogen werden, die die Dimension und die Komplexität des Zustandsrasters erhöhen.

Auf diese Weise wird es möglich, daß ein Zustand, der vor Aufnahme eines bedingten Sollwertes als "Gleichgewicht" oder geringes Ungleichgewicht gewertet wurde, danach ein höheres ^{nat} Ungleichgewicht γ , höher z.B. als ein anderer Zustand, der vor Aufnahme des bedingten Sollwertes in das Raster ungleichgewichtiger war (vgl. Bsp. und Abb. 2.1).

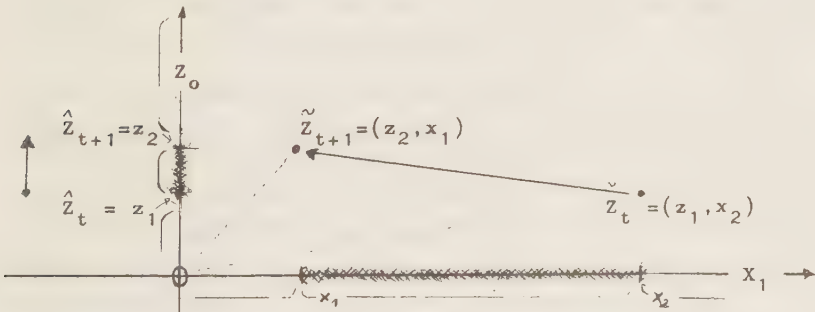


Abb. 2.1

Beispiel: Sei Z_0 um eine Dimension X_1 erweitert. Vor Aufnahme des bedingten Sollwertes X_1 ist der Übergang von $\hat{Z}_t$ nach $\hat{Z}_{t+1}$ mit einer Ungleichgewichtsvergrößerung verbunden. Nach deren Aufnahme bedeutet es hingegen eine Ungleichgewichtsreduktion. (Zur inhaltlichen Bedeutung dieses Ansatzes s. § 3.1.)

Mit Hilfe einer solchen Erweiterung wird es möglich, daß Verhaltensalternativen "verstärkt" werden, die nicht sofort im Bereich der "unbedingten Sollwerte" ($= Z_0$) das Ungleichgewicht reduzieren, dies aber in einem nachfolgenden Takt mit hoher Wahrscheinlichkeit erwarten lassen.

Die vorgenannten Konvergenzsätze sowie deren Erweiterungsmöglichkeiten sind in gleicher Weise für solche mehrstufigen VS anwendbar, wobei es lediglich zu einer veränderten Bewertung des Ungleichgewichts im LS kommt.

Zu klären bleibt der Mechanismus, der ein LS veranlassen sollte, "bedingte Sollwerte" aufzunehmen. Eine Möglichkeit wäre der folgende "Algorithmus":

Wenn nach einer gewissen Versuchshäufigkeit die Ungleichgewichtsreduktion trotz mehrfacher Differenzierung unbefriedigend bleibt, wird ein (mehrstufiges) LS versuchen, Wahrnehmungsdimensionen (d.h. Teilkategorien von X) als bedingte Sollwerte (erweiterte Dimensionen des Zustandsraums) aufzunehmen, deren Rasterung durch die gerade bestehende Rasterung $\mathcal{R}_t$ bestimmt und somit variabel ist. Dabei sind solche Dimensionen mit hoher Wahrscheinlichkeit auszuwählen, die bei dem vorliegenden Zustand in der Vergangenheit zu einer befriedigenden Ungleichgewichtsreduktion geführt haben.

Eine Präzisierung des hier unscharf formulierten Algorithmus ist u.E. nach ohne prinzipielle Schwierigkeiten möglich, würde aber den Umfang der vorliegenden Arbeit sprengen. Andere interessante Bereiche werden angesprochen, wenn man versucht, verschiedene Lernautomaten miteinander bezüglich ihrer Konvergenzeigenschaften oder Konvergenzgeschwindigkeiten zu vergleichen. Nicht zuletzt muß für die Anwendung auf reale LS die Existenz absorbierender Punkte berücksichtigt werden, bei deren Erreichen das LS partiell oder global zerstört wird. Einige der speziell zur Erklärung psychologischer Lerntheorien wichtige Überlegungen werden heuristisch in Kap. III besprochen.

Kapitel III

ERLÄUTERUNG DES FORMALEN MODELLS

3.1 Analyse der Lernphänomene

Wie viele der psychologischen Termini, ist der Begriff des "Lernens" aus der Alltagssprache übernommen worden. Er umfaßt sehr verschiedenartige Phänomene der Verhaltensänderung, die nicht genau beschreibbar, sondern nur über Situations- und Verhaltensbeobachtung von Individuen zu erschließen ist. Die Zahl der Definitionen des Lernens ist beträchtlich, ohne das eine der Definitionen eine eindeutige Allgemeingültigkeit erlangt hätte.

Da die Begriffsbildung "Lernen" wie jedes Konstrukt seine Bedeutung letzten Endes aus den theoretischen Konzepten gewinnt, die mit seiner Hilfe aufgestellt werden, ist es erforderlich, nach den angestrebten Zielen und den zu lösenden Problemen zu fragen. Als zugrundeliegende formale Elemente zur Beschreibung des Lernprozesses können wir zunächst unterscheiden:

- ein lernendes Individuum, charakterisiert durch einen Zustand z (aus einer nicht genau bekannten Zustandsmenge Z)
- eine Stimulussituation x (aus einer grob umrissenen Stimulusmenge X)
- eine Reaktion y (aus einer nur unvollständig präzisierbaren Verhaltensmenge Y).

Lernen im Zeitintervall (t_1, t_2) impliziert in der Regel, daß das Individuum auf einen gegebenen Stimulus x zu zwei verschiedenen Zeitpunkten $t_1 < t_2$ zwei verschiedene Reaktionen $y(t_1)$ und $y(t_2)$ zeigt.

Die Festlegung des funktionellen Zusammenhangs zwischen x und y , sowie speziell dessen Veränderung (also der Lernprozeß) bereitet vor allem deshalb große Schwierig-

keiten, weil der Zustand $z(t_1)$ ebenso wie der Zustand $z(t_2)$ als Realisierungen in einer "black box" aufzufassen sind, die wir im einzelnen nicht kennen und allenfalls mit Hilfe von mehr oder weniger plausiblen Annahmen oder deren Ableitungen erschließen können. (vgl. Abb. 3.1)

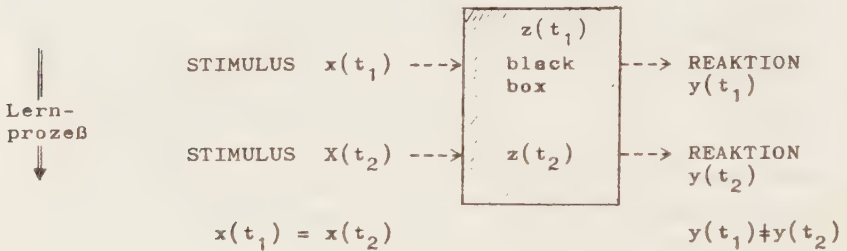


Abb. 3.1

Die dazu notwendige Präzisierung erfordert die Aufstellung einer Theorie des Lernens, die es ermöglicht, Vorhersagen über das Verhalten oder die Verhaltensänderung eines Systems zu geben oder gegebenenfalls manipulativ ("therapeutisch", "pädagogisch") einzugreifen.

Zur Präzisierung der in den Lerntheorien wichtigen Begriffe wie Stimulus (Reiz, Input), Zustand (Struktur) und Verhalten (Reaktion, Output) ist die Angabe der Grundgesamtheit aller zu berücksichtigenden Objekte notwendig, die miteinander in Beziehung gebracht werden sollen.

Nun ist die Zahl möglicher Objekte je nach Anzahl der betrachteten Dimensionen (etwa Größe, Farbe und Form eines bestimmten Stimulus) und nach Feinheit der Kategorien (blau, blau-grau, ultramarin, kobalt hell etc.) unterschiedlich groß. Da die Gesamtzahl möglicher Unterscheidungen unendlich viele Kaggorisierungen in einer großen Zahl von Dimensionen zuläßt, wäre es für ein Lernsystem (=LS) unmöglich, gesetzmäßige Beziehungen zwischen Stimuli und Reaktionen zu entwickeln, sofern es alle unterscheidbaren Aspekte der Wahrnehmung und des Verhaltens bei seiner Zuordnung verwenden

würde. Je nach Situation und nach der Bedürfnisstruktur des LS wird man von unterschiedlicher Kategorisierung der Gesamtheit aller Objekte ausgehen müssen.

Die Schwierigkeit, derartige Zuordnungen in deterministischer Weise vorzunehmen (vgl. etwa die Theorie von HULL in HILGARD/BOWER (1966)), verführt dazu, in eine solche Lerntheorie das Konzept der Wahrscheinlichkeitsbetrachtung einzubauen. Daher ist es sinnvoll, die Variablen Stimulus, Reaktion und Zustand als "Zufällige Variable" (= Stochastische Variable) zu betrachten. Um Verwechslungen zu vermeiden, andererseits eine sparsame Verwendung von Zeichen zu ermöglichen, sind die (zufälligen) Variablen mit einem Häkchen $\wedge$ versehen (z.B. mit $\hat{X}$ der Stimulus), die Grundgesamtheit ihrer Realisierungsmöglichkeiten mit dem gleichen Zeichen, aber ohne Häkchen (z.B. X für den Stimulusraum) und die zugehörigen Kleinbuchstaben (z.B. x) für die tatsächlichen Realisierungen.

3.1.1 Das Lernsystem als Automat

Individuen, deren Verhalten Lernprozessen unterliegt oder unterliegen kann, seien im folgenden als Lernsysteme (LS) bezeichnet. Von System sprechen wir, weil uns nur die strukturellen Beziehungen zwischen Individuum und Umwelt interessieren, und wir nicht danach fragen wollen, ob es sich bei dem System um ein biologisches Wesen oder ein technisches Objekt handelt. Im Rahmen der Automatentheorie wurden bereits zahlreiche spezielle Fragestellungen im Zusammenhang mit Lernprozessen untersucht, insbesondere im Bereich der Wahrnehmungserkennung (vgl. MENDEL/FU, 1971). Wenn wir ein LS als einen speziellen Automaten (Moore-Automaten) einführen, erleichtert das die formale Beschreibung zahlreicher lerntheoretischer Fragen.

Charakteristisch für ein LS ist die Reaktion auf die Umwelt, die über Eingabesignale bedeutsam wird, in Form von Ausgangssignalen (oder Verhaltensalternativen) Y_t . Dabei verändert sich der innere Zustand (Struktur) Z_t des LS.

Die Erweiterung des Modells gegenüber den meisten klassischen, insbesondere den SR-Modellen durch die Einbeziehung des Zustands Z_t des Systems erscheint uns wesentlich, weil erst auf diese Weise Unterschiede zwischen Lernverläufen befriedigend erklärt werden können.

Gemäß dem Modellansatz in Kap. I erweist es sich als ausreichend und zweckmäßig, zur Charakterisierung des Systems im Zeitkontinuum von einer Menge T diskreter einander nachfolgender Zeitpunkte (Zeitklassen) auszugehen, so daß Veränderungen der Stimuli, der Zustands- und der Verhaltenskategorien nur zwischen zwei aufeinanderfolgenden Zeitpunkten t und $t+1$ auftreten können.

Aufgrund der genannten Annahme ist ein LS in seinem dynamischen Aspekten vollständig durch die Angabe der Ereignisse $\hat{E}_t = (\hat{X}_t, \hat{Z}_t, \hat{Y}_t)$, $t \in T$ bestimmt. Ein "Ereignis" in der von uns im folgenden gebrauchten Bedeutung hat also nur im Zusammenhang mit dem angegebenen LS einen Sinn.

In gleicher Weise definieren wir als die Situation $S_t = (\hat{X}_t, \hat{Z}_t)$ das Paar "Stimulus - Zustand" zum Zeitpunkt t . Folglich sind die Schreibweisen $\hat{E}_t$ und $(\hat{S}_t, \hat{Y}_t)$ für ein bestimmtes Ereignis äquivalent.

Von einem "vernünftigen" Verhaltenssystem, d.h. einem Automaten, der als Folge eines bestimmten Eingangssignals $\hat{X}_t$ und eines best. Zustandes $\hat{Z}_t$ ein bestimmtes Ausgangssignal $\hat{Y}_t$ abgibt, erwartet man, daß eine Gesetzmäßigkeit zwischen der Situation (=Eingangssignal + Zustand) einerseits und dem nachfolgenden Verhalten andererseits besteht. LS sollten darüberhinaus in der Lage sein, diesem Zusammenhang aufgrund von Erfahrungen so verändern zu können, daß er bestimmten vorgegebenen Kriterien genügt.

Dieser Zusammenhang zwischen Situation und Verhalten läßt sich durch eine (im vorliegenden Ansatz) stochastische Abbildung wiedergeben, die man in der Automatentheorie als Ergebnisfunktion bezeichnet, formal:

$$[3.1] \quad g_t : S_{(t)} \dashrightarrow Y_{(t)}$$

Bei Lernsystemen muß diese Abb. g_t ($t \in T$) von den vorausgegangenen Erfahrungen, d.h. den Ereignissen $\dot{E}_\tau$ ($\tau < t$) abhängig sein.

Unter den Verhaltenssystemen (VS), die keine Lernsysteme sind, interessieren uns im folgenden die Homöostatischen Systeme (=HS). Bei ihnen bleibt zwar die Abbildung g_t für alle $t \in T$ konstant, sie ist aber so gewählt, daß für einen bestimmten Teilbereich von Umweltsituationen ein nach einem vorgegebenen Kriterium optimales Verhalten als Reaktion auf einen bestimmten Stimulus folgt.

Bei einem Lernsystem ändert sich die Abb. g_t als Funktion der Zeit, jedoch nach einem fest vorgegebenen Kriterium. Auf diese Weise wird die Gesamtheit der bedingten Verhaltensweisen zu einem stochastischen Prozeß (g_t , $t \in T$) und g_t als Funktion der Zeit eine stochastische Funktion (vgl. DOOB, 1965).

Da jedoch das Verhalten eines Systems nicht nur vom Stimulus, sondern auch von dem Zustand des Systems abhängig ist, und jene Beziehung bei Verhaltenssystemen nicht unbedingt zeitlich konstant bleibt, ist zur Beschreibung des Automaten die Kenntnis einer weiteren Abbildung f_t nötig, die den Zusammenhang zwischen Zustand Z_t und dem vorausgehenden Ereignis angibt, formal:

$$[3.2] \quad f_t : E_{(t-1)} \dashrightarrow Z_{(t)}$$

Damit läßt sich ein VS und daher auch ein LS als ein spezieller stochastischer Moore-Automat charakterisieren, wie dies aus der Automatentheorie bekannt ist. (vgl. GLUSKOV, 1963).

Als einer der ersten Versuche, biologische Individuen auf diese Weise zu charakterisieren, kann die Turing-Maschine gelten, die von TURING (1937) konzipiert wurde, als er das Problem untersuchte, ob es logisch möglich sei, ein

System zu entwerfen, das sich selbst reproduzieren kann. Mit seiner "Maschine" lieferte er dazu den konstruktiven Beweis. Die dort begonnenen Überlegungen wurden in verschiedenen, vor allem automatentheoretischen Überlegungen weitergeführt, speziell von VON NEUMANN (1951) und haben zum Teil mittlerweile Eingang in die allgemeine Automatentheorie gefunden. Auch wir haben versucht, Gedankengänge von TURING zu berücksichtigen, wenn auch einem etwas anderen Zusammenhang.

3.1.2 Der Lernprozeß

Ein Automat läßt sich erst dann als ein allgemeines LS darstellen, wenn die verhaltenssteuernde Abbildung g_t eine Funktion der Vorgeschichte ist, d.h. der Automat muß die Erfahrungen der Vergangenheit oder zumindest eines Teils davon zum Zwecke einer Verhaltensmodifikation berücksichtigen können. (vgl. FEICHTINGER, 1970).

Durch die Reduktion des Modells auf diskrete Zeitpunkte müssen natürlich einige der in einem Zeitintervall ankommenden Stimuli, bestehende Zustände und Verhaltensweisen identifiziert werden. Aus formalen Darstellungsgründen erschien es uns zweckmäßig, Stimulus und unmittelbar nachfolgende Reaktion dem gleichen Takt (Zeitpunkt) t zuzuordnen, die über den nächsten Stimulus eventuell wirksam werdende Reaktion der Umgebung auf den nachfolgenden Zustand hingegen dem nachfolgenden Takt $t+1$. Dieser Ansatz schließt nicht die Möglichkeit aus, daß bestimmte Umgebungsreize erst zu einem späteren Zeitpunkt für das Verhalten relevant werden.

Bei der Suche nach Kriterien, denen der Lernprozeß genügen sollte oder könnte, fanden wir in den klassischen Lerntheorien zahlreiche Ansätze, jedoch keine geschlossen verwertbare Theorie (vgl. z.B. KAMINSKI, 1970; MELTON, 1964; oder HILGARD/BOWER, 1971). Daher waren wir gezwungen, eine eigene Zusammenstellung von Postulaten zu versuchen, die einerseits

mathematisch genügend präzisierbar, andererseits psychologisch hinreichend global erschienen. Wir können nicht hoffen, daß die ausgewählten Grundbedingungen vollständig oder wenigstens die zweckmäßigsten sind. Wir haben jedoch auf diese Weise ein uns vielseitig verwendbares Modell konzipiert, das zumindest zu den uns bekannten lerntheoretischen Fragen Stellung beziehen kann.

3.1.3 Das homöostatische Prinzip

Die meisten bisher diskutierten mathematischen Lerntheorien haben den Lernprozeß als einen stochastischen Prozeß aufgefaßt (vgl. BUSH/MOSTELLER, 1955; ESTES, 1964; FEICHTINGER, 1970; HILGARD/BOWER, 1971; NORMAN, 1972). Vor allem die Markoff'schen Prozesse als spezielle stochastische Prozesse sind mit einigem Erfolg bei der Formulierung von Lerntheorien verwendet worden. (s. auch LUCE et al., 1965; STERNBERG, 1963; STEWART, 1967; SUPPES, 1969). Leider erwiesen sich die schon intensiv erforschten Markoff'schen Prozesse für das von uns konzipierte Lernmodell als ungeeignet, weil für ihre Anwendung zu spezielle Forderungen an die Umgebungsreaktionen gestellt werden. Lediglich die noch verhältnismäßig wenig differenzierte Theorie der Semimartingale konnte bei dem Nachweis der Konvergenzeigenschaften nutzbringend verwendet werden. (vgl. Kap. II).

Andererseits zeigt das Studium stochastischer Prozesse, daß die Untersuchung ihrer Eigenschaften umso schwieriger und die Aussagemöglichkeiten über den Verlauf der durch sie bedingten Entwicklung umso geringer wird, in je größerem Maße die Vorgeschichte, d.h. die Vergangenheit des Prozesses zur Erklärung der stochastischen Funktionen herangezogen werden muß. Daher haben wir versucht, als Basis der Lernprozesse möglichst einfache Kriterien zugrunde zu legen. Aus der Kybernetik entnahmen wir das HOMÖOSTATISCHE PRINZIP als Grundlage der von einem Individuum angestrebten Verhaltenstendenz (vgl. CUBE, 1968; FLECHTNER, 1967; FRANK, 1969; WALTER, 1953; WIENER, 1968; YOVITTS/CAMERON, 1960).

Obwohl keine der uns bekannten Lerntheorien auf dieses Prinzip als explizite Erklärungsbasis zurückgreift, erscheint es uns besonders attraktiv, weil es ermöglicht, von einer externen Institution abzusehen, die ein Verhalten des LS belohnt oder bestraft. Damit aber ein LS die Angemessenheit einer Reaktion auf eine gegebene Situation feststellen und gegebenenfalls ^{adäq.} modifizieren kann, ist die Abbildung des Ereignisses $\hat{E}_t$ auf eine Bewertungsmenge (gut - schlecht) erforderlich. Diese Bewertung kann sich sinnvollerweise erst auf eine nachfolgende Situation beziehen, und zwar aufgrund des vom LS dann erreichten Zustands. Ein solcher Bezug auf die nachfolgende Situation verlangt aber das Vorhandensein einer Rückkopplung. Beim Moore-Automaten wird dies durch die Abb. f_t (vgl. [3.2]) gegeben.

Homöostatische Systeme (=HS) sind solche VS, die im Falle eines durch die Umwelt oder durch eigenes Verhalten bedingten Ungleichgewichts in ihrem Zustandssystem (d.h. $\|Z_t\| \neq 0$) ihr Verhalten so zu steuern, daß das Gleichgewicht wieder erreicht wird. Kybernetisch gesprochen muß ein im System vorgegebener Sollwert (oder ein Cluster von Sollwerten) durch den Istwert (bzw. Cluster von Istwerten) approximiert werden, wenn letzterer zunächst vom Sollwert abweicht.

So prüft beispielsweise ein Thermostat dauernd die Abweichung der Isttemperatur von der Solltemperatur. Liegt die Abweichung (etwa durch winterliche Außentemperaturen bedingt) über einem Schwellenwert, dann wird das Verhalten verändert ("heizen"), bis die Approximation an den Sollwert wieder genügend gut ist, mithin das Gleichgewicht in den berücksichtigten Dimensionen wieder hergestellt ist.

Die Funktionsfähigkeit eines HS ist an zwei Bedingungen geknüpft. Erstens muß die Möglichkeit der Kompensation durch das gegebene Verhaltensmuster bestehen.

Die üblichen thermostatischen Einrichtungen können das Temperaturungleichgewicht nicht beheben, wenn es aufgrund sommerlicher Temperaturen zu heiß wird, da sie nur über die Verhaltensalternativen: Heizen/ Nicht heizen verfügen.

Zweitens muß das HS die nachfolgende Situation als Folge des eigenen Verhalten hinreichend gut abschätzen können.

Ein Klimasystem, das zwar über die notwendigen Verhaltensalternativen verfügt (kühlen / heizen / keine Aktivität), kann nicht homöostatisch sein, wenn es nicht jeweils die angemessenen Verhaltensweisen einsetzt, d.h. wenn es nicht bei einer Sollwertunterschreitung den Wärmemechanismus und bei einer Sollwertüberschreitung des Kühlmechanismus in Gang setzt.

Im Falle eines Klimaregulators sieht die Konstruktion eines global wirksamen HS noch relativ einfach aus, zumal sich die Umweltbedingungen in der Regel nur verhältnismäßig langsam verändern und die Vorwegnahme der Konsequenzen des eigenen Verhaltens klar ist. In Fällen, in denen mehr Alternativen zur Verfügung stehen oder die System-Umwelt-Beziehung sensibler als im vorliegenden Fall ist, sind die zu lösenden Probleme bei weitem nicht mehr so einfach.

Betrachten wir als Grundaufgabe biologischer LS die Sicherung ihrer Existenz, dann sind die Homöostatischen Systeme genau die Systeme, die in den jeweils erfaßten Dimensionen die Sicherung der Existenz durch fortgesetzte Annäherung an den Gleichgewichtsbereich ermöglichen (Bsp. Blutzuckerspiegel im Blutkreislauf). Lernsysteme sollten daher so konstruiert sein, daß sie unter hinreichend stabilen Beziehungen zwischen System und Umgebung das Verhalten von HS approximieren lernen, falls dies mit den verfügbaren Verhaltensalternativen überhaupt möglich ist. Dies ist der Grund, warum wir uns in Kap. II vor allem mit der Untersuchung beschäftigt haben, wann eine solche Approximation an ein HS erwartet werden kann.

3.1.4 Das Reinforcementprinzip

Fast allen psychologischen Lerntheorien, selbst den mentalistischen liegt - wenn auch in verschiedener Formulierung und verschiedener Präzisierung - das empirisch evident erscheinende Reinforcementprinzip zugrunde:

Die Belohnung einer bestimmten Verhaltensalternative als Reaktion auf eine gegebene Situation (Stimulus und Zustand !) erhöht die spätere Wiederholungswahrscheinlichkeit der betreffenden Alternative unter gleichen oder ähnlichen Bedingungen. Bestrafung verringert die Wahrscheinlichkeit entsprechend. (HILGARD/BOWER, 1971).

In dem von uns entwickelten Modell wird gemäß 3.1.3 Belohnung mit Verringerung des Ungleichgewichts gegenüber dem letzten Zustand identifiziert, Bestrafung mit Vergrößerung des Ungleichgewichts. Dadurch ist es nicht mehr nötig, extern einen Stimulus als Belohnung oder Bestrafung zu bezeichnen und daraus die Verstärkung abzuschätzen, sondern das Verhalten wird von dem LS selbst intern als "gut" empfunden, wenn es sich als homöostatisch erweist. Konstantes Ungleichgewicht entspricht der neutralen oder irrelevanten Konsequenz. Mit diesem Ansatz können auch komplizierte Verstärkungsvorgänge wie z.B. differenzielle Verstärkung erklärt werden. (vgl. § 4.4).

Von einem brauchbaren LS erwarten wir, daß die bedingten Wahrscheinlichkeiten von Verhaltensweisen $\mu_t(y_k | s_j)$, die bei gleichen oder ähnlichen Randbedingungen wesentlich häufiger und/oder instensiver als die übrigen belohnt werden, schließlich die Grenzwahrscheinlichkeit

$$[3.3] \quad \lim_{t \rightarrow \infty} \mu_t(y_k | s_j) = 1$$

approximieren, hingegen Verhaltensweisen, die entsprechend bestraft oder am wenigsten belohnt werden, zur Grenzwahrscheinlichkeit 0 hin konvergieren. Wenn aber Belohnung mit

Ungleichgewichtsreduktion gleichzusetzen ist, dann entspricht der zu erwartende Lernprozeß einer Approximation an homöostatisches Verhalten.

3.1.5 Das Differenzierungs- und Ökonomieprinzip

Allen uns bekannten streng mathematisch formulierten Lerntheorien ist das Konzept gemeinsam, daß der zugrundeliegende Merkmalsraum konstant gelassen wird, d.h. die Kategorien der Stimuli und des Verhaltens werden von vorneherein festgelegt und nur die Zuordnungswahrscheinlichkeiten werden im Rahmen des Lernprozesses verändert (vgl. BUSH/MOSTELLER, 1955; ESTES, 1964; FEICHTINGER, 1970; MENZEL, 1971; NORMAN, 1972). Jedoch wurden im Rahmen wahrnehmungstheoretischer Überlegungen bereits Konzepte erstellt, die sich mit dem Problem der angemessenen Identifizierung von Stimuli befassen (vgl. BLOCK, 1963; oder MENDEL/FU, 1970).

Als Grundlage der Verhaltensentscheidung (die bei den bisher untersuchten Problemen eine Zuordnungsentscheidung ist), dient die Darstellung der ankommenden Stimuli $\hat{X}_t$ im Hinblick auf die Bedeutsamkeit für das System (vgl. FU, 1968). Dabei lassen sich zwei Arten von Informationsreduktion trennen. Die Begrenztheit sensorischer Medien verringert grundsätzlich die Darstellbarkeit des Umwelteinflusses (= Objektraum) auf bestimmte Dimensionen und bestimmte Genauigkeiten (= Darstellungsraum). Darüberhinaus sind in einer bestimmten Situation $\hat{S}_t = (\hat{X}_t, \hat{Z}_t)$ nie alle sensorisch aufgenommenen Charakteristika eines Stimulus relevant für die Verhaltensentscheidung. Im allgemeinen kann der wahrgenommene Stimulus X''_t auf ein wesentlich größeres Kategoriensystem ohne Schaden abgebildet werden. (vgl. Abb. 3.2).

Für den hungrigen Menschen sind unter allen möglichen Wahrnehmungsaspekten der Stimuli nur diejenigen interessant, die etwas mit der Befriedigung seines Hungers zu tun haben.

Natürlich können sich durch diese Reduktion von $\hat{X}_t$ über $\hat{X}''_t$ auf $\hat{X}'_t$ Irrtümer bezüglich der Wirkung von $\hat{X}_t$ auf den Zustand $\hat{Z}_{t+1}$ des LS ergeben, aber durch die Bewertung der Verhaltenskonsequenzen ist eine Korrektur dieser "Irrtümer" im Rahmen des Lernprozesses grundsätzlich möglich, sofern eine "erfaßbare" Gesetzmäßigkeit überhaupt besteht und sofern ein Mechanismus vorgesehen ist, die die Wahrnehmungs- und die Verhaltenskategorien verändern kann. (vgl. MENDEL/FU, 1971; STEWART, 1967).

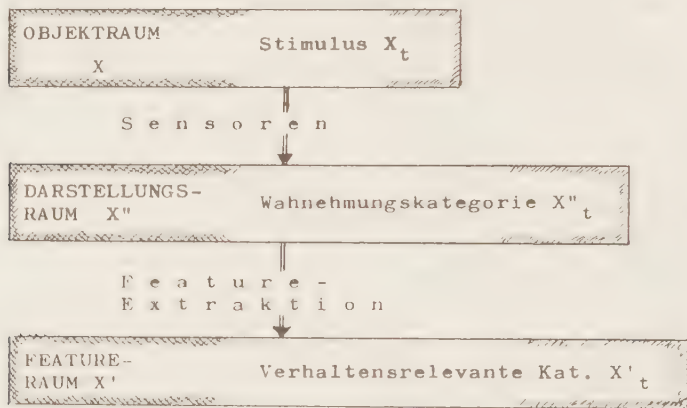


Abb. 3.2

Gilt als Ziel des Lernprozesses nicht nur eine Änderung der Verhaltenswahrscheinlichkeiten auf einen gegebenen Stimulus sondern auch die Ausbildung von neuen Verhaltensweisen oder Wahrnehmungskategorien, dann sind Differenzierungs- und Vereinfachungsmechanismen erforderlich; die Differenzierung zur Ausbildung neuer Alternativen, die Vereinfachung zur Ökonomie der Informationsspeicherung.

Die Einbeziehung differentieller und integrativer Mechanismen verlangt komplexe Überlegungen, die in der Gesamtheit die Möglichkeiten der vorliegenden Arbeit übersteigen. Jedoch haben wir in § 2.3 anhand der verhaltensdifferentiellen LS gezeigt, daß unser allgemeines Modell von Kap.I auch für die Einbeziehung dieses Problems allgemein genug for-

muliert ist. In §2.3 haben wir einen konkreten Weg aufgezeigt, wie selbst bei so komplexen Anforderungen der Lernprozeß mit verhältnismäßig einfachen Mechanismen in der gewünschten Richtung erzielt werden kann. Satz (2.12) zeigt sogar, daß unter relativ allgemeinen Bedingungen die Approximation an ein HS garantiert werden kann. (vgl. Simulationsexperimente mit verh.differentiellen LS in § 4.4). An dieser Stelle sollten jedoch die Überlegungen dargestellt werden, die einen sinnvollen Differenzierungs- und Integrationsmechanismus begründen.

VERHALTENSDIFFERENZIERUNG: Eine Differenzierung der Verhaltensalternativen erscheint uns sinnvoll, wenn für eine gegebene Situation $\hat{S}_t$ der Erwartungswert der Ungleichgewichtsreduktion nicht befriedigend ist. Dann kann das LS nämlich - bei einer Varianz, die nicht verschwindet - unterstellen, daß von den Elementen der Verhaltenskategorien nur ein Teil angemessen ist und somit eine Teilung der Kategorien die Erfolgserwartung verbessern kann.

Die SKINNER'sche Ratte im Käfig, die zunächst nur die die Verhaltenskategorien (stehen/laufen) kennt, lernt bei Belohnung nach zufälligem Berühren einer Taste während des Umherlaufens nach einer gewissen Zeit, die Teilkategorien (Laufen ohne Berühren der Taste/ Laufen mit Berühren der Taste) zu unterscheiden und letztgenannte mit zunehmender Wahrscheinlichkeit auszuwählen.

Da jedoch der Erwartungswert und die Varianz der nachfolgenden Ungleichgewichtsreduktion dem LS nicht bekannt sind, muß es geeignete Statistiken dafür verwenden. In § 2.3 haben wir den Erwartungswert $\sum \hat{f}_t^j(s_j, y_k)$ durch die Stichproben-erwartung

$$[3.4] \quad \left\{ \begin{array}{l} \bar{\mathcal{E}}_{j0}^j(s_j, y_k) = c_{jk} \in [-1, +1] \quad \text{und für } t > 0 \\ \sum \hat{f}_{t+1}^j(s_j, y_k) = 1/n_{1s} \cdot [(n_{1s}-1) \cdot \bar{\mathcal{E}}_{jt}^j(s_j, y_k) + \rho(\hat{S}_{t+1}, \hat{S}_t)] \end{array} \right.$$

mit $n_{1s} \geq 2$ geschätzt. Diese Statistik garantiert, daß bei

einer Varianz der Verh.konsequenzen > 0 eine Verhaltensdifferenzierung durchgeführt wird, sofern gewisse Schwellenwerte überschritten sind, die eine allzuschnelle Differenzierung verhindern.

Wie Satz (2.12) beweist, kann damit die Approximation an ein HS unter verhältnismäßig allgemeinen Bedingungen garantiert werden, selbst wenn dazu eine Verhaltensalternative erforderlich ist, über die das LS zu Anfang noch nicht verfügt, die es aber durch Differenzierung gewinnen kann. (vgl. auch Simulationsexperimente in § 4.4).

WAHRNEHMUNGSDIFFERENZIERUNG: Wahrnehmungsdifferenzierung, d.h. eine Verfeinerung von $\hat{x}_t$ erscheint immer dann sinnvoll zu sein, wenn die Varianz der nachfolgenden Ungleichgewichtsänderung auf eine gegebene Verhaltensalternative und gegebene Situation übermäßig groß ist. Dann kann das LS vermuten, daß die gegebene Reaktion allenfalls für einen Teil der Wahrnehmungskategorie x_j angemessen ist, um eine befriedigende Ungleichgewichtsreduktion zu ermöglichen. Unter bestimmten vernünftig erscheinenden Umständen kann sichergestellt werden, daß auch hier durch einen Differenzierungsalgorithmus die für ein homöostatisches Verhalten notwendige Feinheit des Situationsrasters erreicht wird.

Da auch die Varianz der Ungleichgewichtsänderung dem LS nicht bekannt ist, wählen wir eine entsprechende Statistik wie in [3.4], die Stichprobenvarianz

$$[3.5] \quad \begin{cases} \overline{var} \hat{y}_0(s_j, y_k) = d_{jk}^2 \in [0, 1/4] \quad \text{und für } t > 0 \\ \overline{var} \hat{y}_{t+1}(s_j, y_k) = 1/n_d \cdot [(n_d - 1) \cdot \overline{y}_t \hat{y}_t'(s_j, y_k) + D^2] \end{cases}$$

wobei $n_d \geq 2$, $D = y(\hat{S}_{t+1}, \hat{S}_t) - \bar{x}_t'(s_j, y_k)$ und $\bar{x}_t'(s_j, y_k)$ durch [3.4] definiert ist.

Auch diese Statistik, die der Stichprobenvarianz bei unabhängigen Ereignissen entspricht, konvergiert gegen 0, wenn dies auch für die Varianz der Ungleichgewichtsänderung gilt.

VERHALTENSVEREINFACHUNG: Eine Zusammenfassung von zwei bisher verschiedenen Verhaltenskategorien zu einer Kategorie erscheint dann sinnvoll, wenn die Erwartungswerte der Konsequenzen zweiter Verh.alternativen y_k^i, y_k^u bei ähnlicher Varianz unter einer vorgegebenen Schranke liegt, d.h. formal, wenn gilt:

$$[3.6] \quad \left\{ \begin{array}{l} | \bar{\xi}_{\hat{p}_t}(s_j, y_k^i) - \bar{\xi}_{\hat{p}_t}(s_j, y_k^u) | < \varepsilon' \\ | \bar{u}_{\hat{p}_t}(s_j, y_k^i) - \bar{u}_{\hat{p}_t}(s_j, y_k^u) | < \varepsilon'' \end{array} \right.$$

für gegebene $\varepsilon', \varepsilon'' > 0$. In diesem Fall ist es nämlich gleichgültig, welche der beiden Alternativen das LS als Reaktion auswählt; es kann die Entscheidung dem Zufall überlassen.

WAHRNEHMUNGSVEREINFACHUNG: Entsprechend empfiehlt sich eine Vereinfachung des Wahrnehmungsrasters, wenn die Differenz der Erwartungswerte und Varianzen über alle $y_k \in \hat{Y}$ zu gering ist. Zwei Stimuluskategorien x_j^i, x_j^u werden zu einer neuen Kategorie x_j zusammengefaßt, wenn für gegebene $\delta', \delta'' > 0$ gilt:

$$[3.7] \quad \left\{ \begin{array}{l} \sum_{y_k \in \hat{Y}_t} | \bar{\xi}_{\hat{p}_t}(s_j^i, y_k) - \bar{\xi}_{\hat{p}_t}(s_j^u, y_k) | < \delta' \\ \sum_{y_k \in \hat{Y}_t} | \bar{u}_{\hat{p}_t}(s_j^i, y_k) - \bar{u}_{\hat{p}_t}(s_j^u, y_k) | < \delta'' \end{array} \right.$$

Wie das Beispiel des Verhaltensdiff. LS mit Differenzierungs- und Integrationsmechanismus nach diesem Schema in §4.4 beweist, können diese Algorithmen einfach und wirkungsvoll für die gewünschten Veränderungsprozesse eingesetzt werden.

3.1.6 Transfereigenschaften

Neben dem Reinforcementprinzip ist das auffälligste Kennzeichen realer LS die Fähigkeit zum Transfer gelernter Verhaltensweisen bei ähnlichen Situationen. Noch nicht oder selten aufgetretene Verhaltensalternativen können mit hohen

bedingten Wahrscheinlichkeiten besetzt sein, wenn die betreffende Situation oder das betreffende Verhalten mit häufig gebrauchten Alternativen verwandt ist. Das Problem des Transfers auf verwandte Situationen wurde durch Einführung eines Transferkoeffizienten berücksichtigt, Transfer auf verwandte Verhaltensweisen wird möglich im Konzept der abhängigen Lernsysteme (vgl. §1.2).

3.1.7 System der Sollwerte

Die Basis der Ungleichgewichtsermittlung ist ein System von unbedingten und bedingten Sollwerten. Zu den unbedingten Sollwerten gehören die Dimensionen, die den biologisch funktionsfähigen Ablauf im Organismus des betreffenden LS garantieren. Zu den bedingten Sollwerten, die in unterschiedlich leichter Weise neu aufgenommen oder wieder abgegeben werden können, gehören jene Kategorien (Dimensionen des Zustands bzw. der Wahrnehmung), deren Gleichgewichtspunkt so geartet ist, daß sein Erreichen die Wahrscheinlichkeit wesentlicher Ungleichgewichtsreduktionen maximiert (bzw. die Wahrsch. wesentl. Ungleichgewichtssituationen minimiert).

Die Istapproximation des bedingten Sollwerts:
Nahrungsvorratshaltung minimiert das Risiko,
bei auftretendem Ungleichgewicht des Blutzuckerspiegels nicht in der Lage zu sein, das Ungleichgewicht zu korrigieren, mithin zu verhungern.

Auf diese Weise wird es möglich, daß ein LS Verhaltensweisen durchführt, die zwar nicht sofort zu einer Ungleichgewichtsreduktion führen, insgesamt aber die günstigen Konsequenzen haben. In §2.4 bei der Diskussion mehrstufiger LS wird der formale Apparat, der dazu im Rahmen unserer Modellbildung notwendig ist, kurz diskutiert, ohne erschöpfend sein zu können. Wir möchten an dieser Stelle dazu noch ein bekanntes Tierexperiment als Beispiel zur Verdeutlichung auswählen:

Sei T ein Tier (=LS), bei dem wir folgende vier Zustandskategorien in dem eindimensionalen Zustandsraum Z_0 unterscheiden:

$z_1 :=$ "kein Hunger", d.h. $\|z_1\| = 0$ ¹⁾

$z_2 :=$ "mäßiger Hunger", d.h. $\|z_2\| = 3$

$z_3 :=$ "viel Hunger", d.h. $\|z_3\| = 4$

$z_4 :=$ "s.viel Hunger", d.h. $\|z_4\| = 5$

Um das Problem zu vereinfachen, nehmen wir an, daß keine zusätzlichen Störgrößen auftreten, d.h. das die nachfolgende Situation nur von der vorausgehenden Verhaltensalternative abhängig ist, formal:

$$[3.8] \quad \alpha_{t+1}(s_j | e_i) = \alpha_{t+1}((x_h, z_j) | (x_1, z_i, y_k)) \\ = 0 \text{ oder } 1 \quad \text{für alle } e_i \in \overset{\circ}{\mathcal{E}}, s_j \in \overset{\circ}{\mathcal{S}}_{t+1}$$

Es sei in unserem Beispiel nun folgende Situation gegeben: Das Tier T befindet sich stark hungrig ($\hat{z}_1 = z_3$) hinter einem Zaun, vor dem Futter verstreut liegt. Wenn T den Kopf durch eine der Zaunlücken steckt ($\hat{y}_1 = y_2$), erreicht es einen Teil des Futters, was seinen Hunger teilweise verringern ($\hat{z}_2 = z_2$), aber nicht vollständig befriedigen würde. Um das dafür notwendige Futter zu bekommen, muß T erst um einen Zaun laufen, wobei das Futter zwischendurch aus dem Wahrnehmungsfeld gerät. ($\hat{y}_1 = y_3$) und der Hunger womöglich noch ärger wird ($\hat{z}_2 = z_4$). Macht es das aber, kann es anschließend vor dem Zaun zum Futter laufen und seinen Hunger vollständig stillen. ($\hat{y}_2 = y_4$) ($\hat{z}_3 = z_1$). (vgl. Abb. 3.3).

1) $\|z\|$ bedeutet die Norm von z in einem metrischen Raum. $\|z\| = 0$ heißt somit, daß dessen Nullpunkt erreicht ist, den wir als Gleichgewichtspunkt interpretieren wollen. $\|z\| > 0$ bedeutet dann das Bestehen eines Ungleichgewichts. Die Ungleichgewichtsänderung wird durch die Differenz zweier zeitlich nachfolgender Zustandsnormen ermittelt und mit $\dot{f}_t = f(\hat{z}_{t+1}, \hat{z}_t) = \|\hat{z}_{t+1}\| - \|\hat{z}_t\|$ bezeichnet.

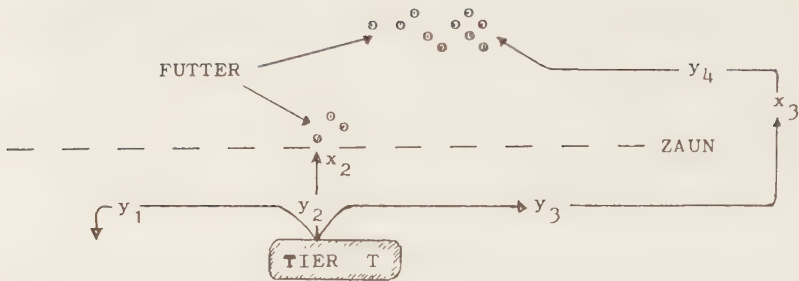


Abb. 3.3

Wichtig zu unterscheiden sind hier folgende Wahrnehmungskategorien:

- x_1 : T sieht Futter, Zaun ist dazwischen
- x_2 : Ein Teil des Futters wird T unmittelbar zugänglich, aber der größere Teil nicht.
- x_3 : Das gesamte Futter wird nun für T zugänglich

Bei den Verhaltensalternativen müssen unterschieden werden:

- y_1 : T begeht eine bezüglich des Futters irrelevante Handlung
- y_2 : T steckt den Kopf durch den Zaun und frisst einen Teil des Futters.
- y_3 : T läuft an das Zaunende (ohne unmittelbaren Gewinn)
- y_4 : T läuft vor den Zaun und frisst das gesamte Futter.

Die Ausgangssituation ist also $S_1 = (x_1, z_3)$, die Endsituation je nach den Verhaltensweisen $S_3 = (., z_2)$, $S_3' = (., z_1)$ oder $S_3'' = (., z_4)$.

Ein LS, das nur unbedingte Sollwerte besitzt, kann nicht lernen, die Verhaltenskette (y_3, y_4) zu wählen, um die Situation S_3' zu erreichen, denn das Bestreben zur Ungleichgewichtsminimierung wird es das Verhalten y_2 als bestes zur Hungerreduktion erlernen lassen. Besitzt T jedoch neben dem unbedingten Sollwert "Hunger ist befriedigt" noch den bedingten Sollwert "Freier und direkter Zugang zu ausreichend viel Futter", dann wird es bei hinreichend hoher Gewichtung des bedingten Sollwerts die Alternative y_3 der Alternative y_2

vorziehen, weil sie im erweiterten Sollwertsystem zu einer stärkeren Ungleichgewichtsreduktion führt (vgl. Abb. 3.4). Beide Fälle können wir in realen LS-Experimenten finden. Das Huhn denkt nur direkt und kann daher nicht um den Zaun laufen, wenn es dabei das Futter aus den Augen verliert, der etwas intelligentere Hund dagegen ist zu einer solchen Entscheidung aufgrund seiner Lernprozesse bereits in der Lage.

Die Formalisierung im Rahmen der Modelldarstellung, wie sie in § 2.4 skizziert wurde, könnte bei diesem Beispiel etwa so aussehen: Der Zustandsraum Z_0 wird um die Wahrnehmungsdimension X (mit den Kategorien x_1, x_2, x_3) erweitert, auf dem dann das neue Zustandsraster gebildet wird. (vgl. Abb. 3.4).

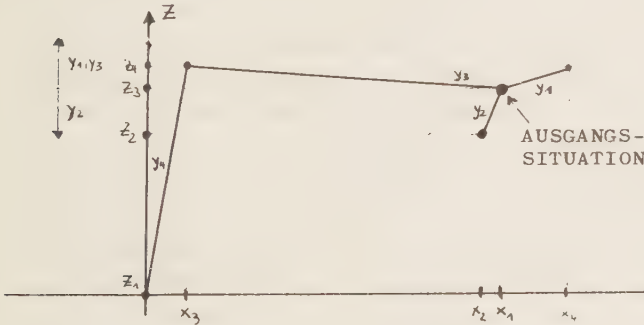


Abb. 3.4

x_3 läßt mit hoher Wahrscheinlichkeit eine vollständige Ungleichgewichtsreduktion erwarten, x_2 eine teilweise Red. und x_1 keine sofortige Reduktion. Daher kommt die Rangfolge der Kategorien auf der X -Dimension.

y_2 , das im System $\tilde{y}_1$ sogar noch eine Vergrößerung des Ungleichgewichts bewirkt, kann im System $\tilde{y}_2$ eine deutliche Ungleichgewichtsverringerung erreichen, so daß es im System $\tilde{y}_3$ der Alternative y_2 vorgezogen wird, nicht aber im System $\tilde{y}_4$.

Auch bei Menschen ist diese Substitution von Sollwerten deutlich. In wirtschaftlichen Normalzeiten bleiben die Abweichungen von den unbedingten Sollwerten gering, wenn bestimmte bedingte Sollwerte (z.B. Geldbesitz) erfüllt sind, so daß ledig-

lich letztere zur Maxime des Handelns werden. Erst wenn die unbedingten Sollwerte extrem weit von den Istwerten erfüllt sind ("Durst in der Wüste" o.ä.) wird die Erfüllung bedingter Sollwerte nachrangig.

Die angegebenen Beispiele mögen etwas einfach sein; sie charakterisieren aber die Möglichkeiten, die mit der Aufnahme bedingter Sollwerte gegeben sind, ohne zu unübersichtlich zu sein. Dieses Prinzip läßt sich (vor allem in noch höherstufigen Formen = sukzessive Substitution) auf auf sehr viel interessantere Problemstellungen anwenden, muß allerdings wegen des Aufwandes an dieser Stelle unterbleiben.

*

3.2 Diskussion der Grundannahmen für ein LS

Zur Definition des VS (1.3): Die atomare Zerlegung des Ereignisraums als Basis der zu untersuchenden bedingten Wahrscheinlichkeiten, die die für uns wichtigen Aspekte der stochastischen Prozesse $(f_t, t \in T)$ und $(g_t, t \in T)$ charakterisieren, erfordern die für die Anwendung durchaus sinnvolle Einschränkung, daß jede atomare Zerlegung Verfeinerung oder Vergröberung des Vorgängers ist.

Die in [1.1] definierte Funktion $y(\dots)$, die als Ungleichgewichtsänderung bezeichnet wird, dient als Basis des Korrekturmechanismus der bedingten Wahrscheinlichkeiten im LS. Die in [1.2] vorgenommene Normdefinition über das Infimum statt wie üblicherweise durch das Supremum ist aus inhaltlichen Gründen sinnvoll. Für biologische und technische Systeme gibt es Schwellenwerte und Toleranzgrenzen. Unterhalb der Schwellenwerte wird eine Störung bzw. eine Aktivität nicht mehr als eine solche registriert (=Nullbereich). Auf der andern Seite wird das System in der Regel in den Extrembereichen, deren untere Grenze (=Infimum) durch die Toleranzwerte festgelegt sind, zerstört, gleichgültig, in welchem Maße sich das Ungleichgewicht $\|\hat{Z}_t\|$ weiter vergrößert. Außerdem ist es zumindest in Teilsystemen der Fall, daß eine Ungleichgewichtsreduktion nur in bestimmten Teilbereichen von Z angestrebt wird und die Abweichung von 0 in den übrigen Dimensionen vernachlässigt werden kann.

Die in [1.4] beschriebenen Abbildungen ermöglichen die algorithmische Realisierung von Differenzierungs- und Verfeinerungsmechanismen, sowie die Veränderung des Sollwertsystems und zwar in Abhängigkeit von den gemachten Erfahrungen (vgl. auch die Computersimulation in §4.4).

[1.5] läßt die Möglichkeit beliebiger Manipulationen seitens der Umgebung offen. Die von FEICHTINGER (1970), IOSIFESCU/THEODORESCU (1969) und NORMAN (1972) gemachten Annahmen einer Umwelt mit konstanten Reaktionswahrscheinlichkeiten

x_t sind für ein allgemeines Lernmodell zu eng. Jedoch müssen die systematischen Änderungen in einem bezüglich $\|\hat{z}_t\|$ kritischen Bereich hinreichend selten sein und die Varianz der zufälligen Störungen in diesem Bereich hinreichend gering, wie die Sätze von Kap. II beweisen.

[1.6] gibt die Abhängigkeit des Lernprozesses von dem letzten Takt ereignis und nur von diesem wieder. Dies ist mindestens nötig, wenn $\gamma(.,.)$ als Grundlage des Lernprozesses herangezogen werden soll. Über die Aufnahme bedingter Sollwerte können aber indirekt weitere Ereignisse der Vergangenheit berücksichtigt werden.

Zur Definition des HS (1.6): Voraussetzung der Funktionsfähigkeit eines HS ist die Bedingung, daß die durch Reduktion des Ungleichgewichts eingesetzte Verhaltensalternative auch wirklich ungleichgewichtsreduzierend wirkt. Da das Funktionsschema eines HS im "voraus" generell festgelegt wird, müssen die Gesetzmäßigkeiten bekannt sein, die der Umweltreaktion zugrundeliegen. Außerdem darf die Dichte "wesentlicher" unvorhergesehener Störungen nur so groß sein, daß eine Korrektur des durch die Störung bedingten Ungleichgewichts möglich wird, ehe eine neue Störung auftritt.

Die Funktionsfähigkeit des Thermostaten als automatischer Temperaturregler beruht darauf, daß erstens die Gesetzmäßigkeiten der Verhaltensalternativen und ihrer Umweltkonsequenzen sich nicht wesentlich verändern. So würde etwa ein Rohrbruch im Heizwassersystem die Gesetzmäßigkeit zwischen der VA "Heizen" und der Umweltreaktion "Raumtemperatur" entscheidend stören. Zweitens müssen die Umweltstörungen (Kälteeinbruch, Hitzewelle) so selten auftreten, daß das Klimasystem korrigierend wird, ehe eine neue wesentliche Temperaturänderung auftritt.

Zur Definition des LS (1.11): Prinzip der Änderung der Verhaltensalternativwahrscheinlichkeiten μ_t ist die Addition eines sogenannten Lernincrements λ_t . Dies bewirkt im Falle, wo $\psi(\hat{Z}_{t+1}, \hat{Z}_t) < 0$ ist, eine Erhöhung der Wahrscheinlichkeit, für $\psi(\hat{Z}_{t+1}, \hat{Z}_t) > 0$ eine Verringerung. Dank des Algorithmus führt diese Addition im mittleren Bereich zu einer raschen Änderung der Wahrscheinlichkeiten (d.h. wenn $\mu_t \approx 1/2$), hingegen in den Extrembereichen zu einer langsamen Änderung. Da der Algorithmus weiter sicherstellt, daß zu allen Zeiten die Wahrscheinlichkeiten der einzelnen V.alternativen > 0 sind, können systematische Änderungen der Gesetzmäßigkeiten zwischen Verhalten und Zustand irgendwann vom LS registriert werden und ein Umlernen ermöglicht werden.

Die Approximation an ein "erfolgreiches", nicht notwendig an das beste Verhalten wurde in Kap. II nachgewiesen. Diese in gewissem Sinne "Nichtoptimalität" von LS ist der Preis für die größere Variabilität gegenüber veränderten Verhältnissen.

3.3 Konsistenz des Modells mit empirisch gestützten Lernphänomenen

Im folgenden Abschnitt soll skizziert werden, wie sich die empirisch "evidenten" Lernphänomene im Rahmen des hier entwickelten Lernmodells erklären lassen.

Das Reinforcementprinzip wurde unmittelbar in die Konzeption aufgenommen und wird beim LS wirksam durch die Art des Lernincrements, wenn man die Äquivalenz von Belohnung und Ungleichgewichtsreduktion einerseits und Bestrafung und Ungleichgewichtserhöhung andererseits akzeptiert.

Lernen durch Übung ist in unserem Modell evident, da die Zunahme der "belohnten" Verhaltenswahrscheinlichkeiten, wie auch die Computereggebnisse im Anhang beweisen, in der typischen Lernkurve verläuft.

Das klassische Konditionieren läßt sich als die Tendenz zur Vereinfachung der differentiellen Wahrnehmungsstruktur mit Hilfe des Differenzierungs- und Ökonomieprinzips erklären, das ja im allgemeinen Modell des LS berücksichtigt wird.

Das Gesetz der Bereitschaft ("THORNDIKE") wird in unserem Modell dahingehend interpretiert, daß die Reaktion des LS umso günstigere Auswirkungen erwarten läßt, je angemessener die Differenzierungsstruktur des LS bezüglich der vorliegenden Situation ist.

Auch die Extinktion bei fehlender Belohnung oder Bestrafung zunächst konditionierter Reize folgt, weil die Wahrscheinlichkeitszunahme einer Alternative automatisch die Abnahme der übrigen bewirkt und umgekehrt. Ebenso folgt hier die Tatsache, daß in der Regel Belohnung schnelleres Lernen als Bestrafung bewirkt. Sind nämlich mehr als 2 Alternativen vorhanden, dann führt die Belohnung zur vollen Wirkung des Lernincrements für die betreffende Alternative,

während die Bestrafung einer andern Alternative das entsprechende positive Lernincrement auf mehrere Alternativen verteilt. Für die gewünschte Alternative ist sie dann von entsprechend geringerer spezifischer Wirkung. Daneben kommt aber noch ein sozialer Effekt ins Spiel. Bestrafung im sozialen Bereich ist häufig keine reine Ungleichgewichtserhöhung, da es einen "das Ungleichgewicht reduzierenden" Aufmerksamkeits-effekt hat. Sekundäre Verstärkung tritt in dem Moment auf, wo Situationen, die selbst nicht ungleichgewichtig sind, durch Lernprozesse bei der Hierarchiebildung (Metapläne) selbst Sollwertcharakter bekommen. Gleiches gilt für die intermittierende Verstärkung, das verstärkende Element wird höherstufig, und daher die einstufigen Mißerfolge noch nicht notwendig relevant für ein negatives Lernincrement.

Das Gesetz der Prägnanz der Gestaltpsychologen ergibt sich beim KS-Modell wieder unmittelbar aus dem Ökonomieprinzip. Als "Überlegung" könnte man hier den Übergang von Metaplänen zu Subplänen oder umgekehrt bezeichnen, weil es Fälle gibt, bei denen das optimale Level der Reaktionsbestimmung nicht von vorneherein klar und generell keineswegs eindeutig ist.

Die retro- und proaktiven Hemmeffekte sind wahrscheinlich ein Problem der nur eingeschränkt möglichen Wahrnehmung, solange eine Verhaltensentscheidung durchgeführt wird. Unter Umständen wird dann die Speicherung des abgelaufenen Ereignisses teilweise verhindert, ebenso die angemessene Einstellung auf den neuen Stimulus.

Um die Begrenztheit der Aktivitätsleistung im Rahmen des KS-Modells erklären zu können, ein Aspekt, der vor allem bei der Feldtheorie von LEWIN eine Rolle spielt, müßte man zusätzlich zu den statischen Sollwerten einen unbedingten dynamischen Sollwert einbeziehen. Er ist erreicht, wenn eine bestimmte, wahrscheinlich altersabhängige, Aktivität realisiert wird. Seine Erfüllung rangiert jedoch - zumindest innerhalb gewisser Grenzen nach der Erfüllung der unbedingten statischen Sollwerte.

Ein Kleinkind, dessen Bedürfnisse alle befriedigt sind, fängt an, Neugierverhalten zu zeigen (Spieltrieb, Bewegungstrieb). Nur mit Hilfe statischer Sollwerte wäre dieser Effekt nicht zu erklären.

Was LEWIN in seiner Theorie mit Umstrukturierung des Feldes bezeichnet, läßt sich bei uns mit dem Entscheidungsprozeß, d.h. der Auswahl von Meta- und Subplänen interpretieren, die durchaus einen gewissen Zeitraum benötigen können, sowie mit der Bildung bedingter Sollwerte (von LEWIN als Quasiabedürfnis bezeichnet).

Auch der Effekt der Sättigung kann als Abkehr von bedingten Sollwerten aufgrund der geringen Auslastung des Sollwerts "Aktivität" in diesem Zusammenhang interpretiert werden. Entsprechend ist dann das Regressionsverhalten eine Abkehr von bedingten Sollwerten als Folge der Überbeanspruchung der Aktivität oder als Folge eines nicht lösbaren Konflikts.

Interessant ist, daß sich mit diesem Modell sogar die von FREUD schwerpunktmäßig untersuchten Lernvorgänge erstaunlich gut interpretieren lassen. Die zentrale Behauptung der Tendenz zur Spannungsverminderung entspricht bei uns dem zentralen Konzept der Tendenz zur Ungleichgewichtsreduktion. Den Konflikt als wesentliche Basis des Lernens bei FREUD erklären wir damit, daß die verfügbaren Alternativen so beschaffen sind, daß jede der eine Dimension des Ungleichgewichts reduzierenden Alternativen eine andere Dimension des Ungleichgewichts wieder erhöht. LS mit bedingten Sollwerten setzen je nach verfügbaren Plänen unterschiedliche Verhaltensweisen ohne Optimierungsmöglichkeit ein, die von FREUD als Abwehr- oder Verteidigungsmechanismen bezeichnet werden, z.B. Verschiebung, Projektion, Reaktionsbildung, Konversion, Regression oder Sublimierung. Alles dies sind im KS-Modell besondere Formen der Ungleichgewichtsverringering, die als Folge der Veränderung der Hierarchie der Sollwerte zustande kommen.

Kapitel IV

ANWENDUNGSMÖGLICHKEITEN IN DER EMPIRISCHEN PSYCHOLOGIE

4.1 Koppelung von Systemen

Interessante Anwendungsmöglichkeiten ergeben sich, wenn wir nun Systeme von zwei oder mehr VS betrachten, die mit einander in bestimmten Beziehungen stehen. Auch hier können die angeschnittenen Probleme nicht erschöpfend behandelt werden; sie dienen lediglich der Demonstration von Einsatzmöglichkeiten.

Die einfachsten Untersuchungen über psychologische Lernprozesse lassen sich durchführen, wenn ein gekoppeltes System von zwei VS vorliegt, d.h. wenn zwei Systeme so miteinander verbunden sind, daß der Output des einen zum Input des andern wird (vgl. Abb. 4.1). Streng genommen kommen diese Fälle in der Praxis nicht vor, aber für viele Probleme genügt es zur Beschreibung der ablaufenden Lernprozesse.

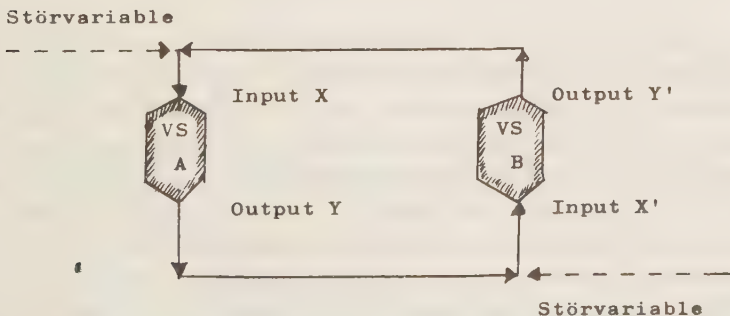


Abb. 4.1

Wir beschränken uns auf diese Zweiersituation, um den Umfang der zu diskutierenden Alternativen nicht zu sehr anwachsen zu lassen. Prinzipiell läßt sich das Modell auf

beliebig viele Systeme in beliebigen Konstellationen anwenden, in denen Lernprozesse ablaufen. Nehmen wir B als ein Lernsystem an, dann sind folgende Alternativen für A von besonderem Interesse:

- A ist ein Homöostatisches System
- A ist selbst ein Lernsystem gleicher oder verschiedener Komplexität
- A ist ein beliebig manipulierbarer Mechanismus, der auf B in einer gesetzmäßigen Weise reagieren, aber auch unabhängige Outputs an B abgeben kann.

Der letztgenannte Fall liegt vor beim vollkommenen Lehrer oder Therapeuten, der ganz nach Belieben auf das LS (Patient, Schüler) eingeht oder unabhängige Steuerungsimpulse abgibt.

Den zweiten Fall finden wir in gewisser Annäherung in den einfachsten Sozial- und Kommunikationssituationen. In diesen Fällen wird das Verhalten der Individuen zueinander weitgehend durch die eigenen Bedürfnisse (Zustand) und das Verhalten des Gegenüber (Output von A = Input von B) gesteuert.

Der erste Fall liegt in geeigneten zeitlichen und örtlichen Ausschnitten der natürlichen Umwelt vor. Die meisten Interaktionen zwischen Individuen und natürlicher Umgebung, aber auch kultureller Umgebung verlaufen in Gesetzmäßigkeiten, denen das Prinzip der Homöostasie, d.h. Tendenz zum Gleichgewicht unterliegt. Ist nun A ein HS, dann gibt es ein geeignetes Situationsraster von A, das quasistabil ist, und bezüglich dessen das System homöostatische Verhaltensweisen einsetzt (vgl. Def.(1.6)). Damit die gesamte Input-Outputfolge der Systemverknüpfung regelnd korrigiert, d.h. konvergiert, muß zunächst einmal B Verhaltensweisen besitzen oder entwickeln, die bei auftretenden Ungleichgewichtssituationen von A für A ein Stimulus sind, für die es eine ungleichgewichtsreduzierende Verhaltensalternative (=VA) besitzt.

In leichter Verallgemeinerung wird im Falle zweier LS lediglich die Existenz von ungleichgewichtsreduzierenden VA gefordert, da sie durch Erfahrungen von beiden LS gelernt werden können.

Für eine zweckmäßige Interaktionssteuerung (pädagogischer oder verhaltenstherapeutischer Art) lassen sich aus der Modellentwicklung unmittelbar einige Forderungen ableiten. Im folgenden wollen wir mit Lernziel (=LZ) eine endliche oder unendliche Kette von Ereignismengen ($E_1, E_2, \dots$) bezeichnen, die aus der Menge aller denkbaren Ereignisketten vorgegeben sei (bezogen auf das LS B). Diese ungewöhnlich erscheinende Benennung entspricht dennoch der üblichen Diktion, wo man ein Lernziel dadurch charakterisiert, daß man auf eine gegebene Situation s_j eine bestimmte Reaktion y_k , zusammen also ein bestimmtes Ereignis $e_i = (s_j, y_k)$ erwartet. Dehnt man diese Bedingung auf mehrere zusammenhängende Ereignisse aus, hat man die a.g. Definition. Ein pädagogisches oder psychologisches Steuerungssystem A sollte so beschaffen sein, daß beim LS B im Zusammenhang für jede der Situationen, die bei ihm auftreten, jenem ein homöostatisches Verhalten möglich ist. Zur Beschleunigung der Konvergenz sollte dieses homöostatische Verhalten beim LS keine zu geringe Wahrscheinlichkeit besitzen.

Diese zunächst rein formalen Forderungen implizieren jedoch für die pädagogische oder therapeutische Praxis bedeutsame Konsequenzen:

- Die für das LS relevanten Lernsituationen müssen ihm möglichst oft geboten werden, d.h. aber nicht nur die Stimuli, sondern auch die relevanten Zustände (!)

Im Rahmen der S-R Theorien stellte sich dieses Problem nicht; dies führte selbst in zahlreichen Anwendungsgebieten zu fast behavioristischen Konsequenzen.

- Um die Konvergenzgeschwindigkeit des Lernprozesses zu erhöhen, ist einer Verstärkung des LS nach einem sich verfeinerndem Schema nötig, denn bei einer bestimmten Wahrscheinlichkeit von $p = 0.5$ ist die Änderungswahrscheinlichkeit (das Lernincrement) am größten und somit der Lernprozeß am schnellsten.
- Die vom Therapeuten oder Lehrer als Belohnung angesehenen nachfolgenden Zustände müssen von dem LS auch wirklich als Belohnung angesehen werden.

Zugleich wird anhand der Modelldiskussion deutlich, welche schwer lösbaren Probleme für Therapeuten und Lehrer auftreten:

- Das LS lernt umso schneller, je seltener bezüglich des Lernziels irrelevante Situationen auftreten. Die Relevanz der Situation läßt sich aber für den Außenstehenden nur grob abschätzen, weil er den Zustand Z des LS nur unvollständig oder gar nicht kennt.
- Eine Konvergenz des Lernprozesses ist nicht möglich, wenn sich das belohnte Verhalten bei ähnlichen Situationen wesentlich verändert. Dies tritt aber oft dann ein, wenn die Therapie- oder Shhulsituation im Gegensatz zur Alltagssituation steht (Elternhaus, Peergruppe etc.). Wichtig, aber nicht leicht zu realisieren, wäre es daher, das LS die Experimentalsituation als einen - vielleicht speziellen - aber doch typischen Fall der allgemeinen Lebensumstände erfahren zu lassen.

4.2. Vorhersageprobleme im pädagogischen Bereich

In der Möglichkeit, Vorhersagen über Lernprozesse und deren Grenzwerte unter bestimmten Anfangsbedingungen zu machen, dürfte ein wichtiges und attraktives Anwendungsgebiet des Modells liegen. Dabei lassen sich verschiedene Fragestellungen unterscheiden:

- Auslase: Gegeben ist das Belegungssystem (= BS = z.B. Lehrer) und das Lernziel (=LZ). Gesucht ist das LS, das das LZ innerhalb einer vorgegebenen Taktzahl erreicht bzw. hinreichend gut approximiert.
- Ausbildung: Gegeben ist das LS und das LZ. Gesucht ist das BS, das in einer vorgegebenen ^{Zeit}spanne das LZ hinreichend gut durch das LS erreichen läßt.
- Platzierung: Gegeben sind mehrere BS, LS und Lernziele. Gesucht ist die Zuordnung, die insgesamt die beste Approximation des zu erreichenden ^{LZ} bei gegebenem BS und LS erwarten läßt.
- Prognose: Gegeben ist das Lehrersystem, das LS und ein Lernziel. Ein Teil des Lernprozesses ist bekannt. Gesucht ist eine Abschätzung der weiteren Ereignisfolge und damit im Zusammenhang die Frage nach der Güte der erreichbaren Approximation an das Lernziel.
- Prognostische Diagnose: Gegeben ist das Lehrersystem und ein Folgenanfang des LS. Gesucht ist die Struktur des LS und die Menge der in einem bestimmten Zeitraum erreichbaren Lernziele.

Auch für dieses Problem können hier nur einige Hinweise gegeben werden. Da das Lernziel Teilmenge des Ereignisraumes oder eines Produkts davon ist, gilt es für einen gegebenen Anfangszustand festzustellen, von welchen Faktoren das Erreichen des Lernziels abhängig ist. Ist das Lernziel ein direkt homöostatisches Verhalten in einer gegebenen Situation, wissen wir bereits nach Kap. 2, unter welchen Bedingungen eine Approximation an das gewünschte Verhalten auftritt. Schwieriger wird die Beurteilung, wenn es sich nur um Teilmengen eines homöostatischen Verhaltens höherer Stufe handelt.

Es gilt dann nämlich festzustellen, ob die notwendigen Ungleichgewichtssituationen als Voraussetzung des Lernens als Folge der vom LS zu bildenden bedingten Sollwerte (Ansehen, Anerkennung, gute Noten u.ä.) im LS auch tatsächlich bestehen.

4.2.1 Auslese

Diese Frage steht in pädagogischen Überlegungen dort im Vordergrund, wo eine bestimmte Fähigkeit des LS als eine nicht wesentlich änderbare Größe angenommen wird; sei es aus prinzipiellen Gründen (Evolutions- und Rffungstheorien) sei es aus Gründen des kurzen möglichen Lernprozesses. In diesem Fall trachtet man danach, für ein gegebenes Lernziel und geg. Lehrersystem den geeignetsten Probanden als Lernsystem zu finden.

Unterstellt man z.B. aufgrund eines streng begrenzten Zeit- . raumes der Lernphase eine bestimmte Grenze der Lernmöglichkeit einer gegebenen Population von LS, dann kann man danach fragen, welche Gruppe von LS bei einer gegebenen Didaktik in der Lage ist, ein vorgegebenes Teilziel zu erreichen.

Notwendige Bedingung für das Erreichen des Lernziels durch das LS ist einmal, daß die in der Didaktik angebotenen Belohnungen für die gewünschten Verhaltensweisen beim LS zu Ungleichgewichtsreduktionen führen, zum anderen aber auch das Bestehen eines bereits hinreichend differenziierten Wahrnehmungs- und Verhaltensrasters oder zumindest die Fähigkeit zu einer raschen Differenzierung in der gewünschten Richtung. Derartige Voraussetzungen werden oft mit Hilfe von Tests oder Vorformen dazu (Noten, Befragung, Eindrucksurteil u.ä.) überprüft.

4.2.4 Prognose

Die Beantwortung der Frage, wann für ein gegebenes Tripel (LS, Lehrersystem, Lernziel) bestimmte Teilziele erreicht werden, setzt nicht nur die Kenntnis über die Anfangsbedingungen der Systeme voraus, sondern auch eine hinreichende Abschätzung der Konvergenzgeschwindigkeit. Für letztere spielt aber neben der Belohnungswahrscheinlichkeit die Güte des Differenzierungs- und gegebenenfalls des Integrationsmechanismus eine wesentliche Rolle. Abschätzungen dafür sind - unter der Annahme bestimmter grundlegender Gesetzmäßigkeiten - durch die Analyse des bisherigen Lernverlaufs des Systems möglich (= Statistisches Problem). Hierbei ist es jedoch analog zur sequentiellen Entscheidung in der Statistik von Nutzen, wenn man durch neue Erfahrungen des Verlaufs eine jeweils neue und damit bessere Schätzung der Weiterentwicklung vornimmt.

Eine weitere Voraussetzung einer guten Prognose ist die Abgrenzung von Lernzielen und die Menge der zulässigen Hilfsmittel.

- Im pädagogischen Bereich impliziert das die Operationalisierung von Lernzielen sowie der Didaktiken; daneben aber vor allem den Vergleich mit den späteren Praxissituationen. Speziell in der Schule hiesse das: Mehr praxisnahe Tätigkeit, nicht nur wegen der Möglichkeit, dem Schüler sicheren Transfer zu ermöglichen, sondern auch, weil auf diese Weise der Schüler durch die Rückkopplung der eigenen Erfahrung aus einer Menge unübersichtlicher Möglichkeiten schneller die für ihn angemessenen herausfinden kann.

Insgesamt gesehen würde das bedeuten, daß eine gute Prognose immer mit einer guten Hilfestellung verknüpft sein sollte, und sich nicht auf einen einzelnen Zeitpunkt beschränken kann.

4.2.5 Prognostische Diagnose

Der letztgenannte Fall ist die Regel in der Schule. Lehrpläne und Lehrer sind als mehr oder weniger festgelegte Didaktiken vorhanden, daneben liegen die ersten oder sogar beträchtlichen Erfahrungen über den Schüler vor, festgehalten durch ausgewählte Ereignisse oder Ereigniskummulierungen (Prüfungen, Noten, Gesamtbeurteilung). Die Frage stellt sich nun entweder allgemein nach der Menge der hier errichtbaren Lernziele in einem vorgegebenen Zeitraum, oder speziell nach der Entscheidung, ob ein bestimmtes Lernziel von einem bestimmten LS in einem vorgegebenen Zeitraum ausreichend approximiert wird.

Voraussetzung einer brauchbaren Diagnose ist die Kenntnis der LS-basis, d.h. der bestehenden Differenzierung von Wahrnehmung und Verhalten, ferner die verfügbaren Differenzierungsmechanismen sowie der unbedingten und der bedingten Sollwerte und deren Abhängigkeit von den Umweltbedingungen (vgl. auch SKOWRONEK, 1970). Hinweise geben Tests, wenn sie dimensions- und strukturelevant sind, ferner die Untersuchung aktuell und latent vorhandener Ungleichgewichtssituationen sowie die Kenntnis der Wirkung bestimmter "Belohnungen" auf eben jenes Ungleichgewicht.

Während die Anwendung von Tests zur Diagnose mit wechselndem Erfolg bereits seit längerer Zeit in der Psychologie und Pädagogik versucht wird, gibt es für die Prüfung der Ungleichgewichtssituationen nur sporadische Ansätze.

- Als ein bekanntes, wenn auch umstrittenes Beispiel, Ungleichgewichtsreduktionen selbstregulierend zuzulassen und damit zu garantieren, kann die Schule in Summerhill gelten (vgl. O'Neill (1971)). Jedoch fehlt bei dieser Variante wiederum der Nachweise, daß mit der gegebenen Methode die gewünschten bzw. die für notwendig erachteten Lernziele auch tatsächlich approximiert werden.

Bestrebungen müßten dahingehen, ein didaktisches Regelsystem zu konzipieren, das sowohl die Reduktion des Ungleichgewichts (Erfüllung, Freude, Lust) im Sinne von zu sichernden Minimalbedingungen kontrolliert, wie auch die Annäherung an das Lernziel (Leistung, Sozialverhalten, Gesundheit u.ä.) sichert. (vgl. auch CUBE, 1968 und YOVITTS/CAMMERON, 1960).

4.3 Anwendungen in der Verhaltenstherapie

Wie die Simulationsbeispiele am Schluß der Arbeit zeigen, können mit Hilfe des Modells Methoden der Verhaltensformung und anderer verhaltenstherapeutischer Maßnahmen treffend mit Hilfe eines Computerprogramms simuliert werden.

Die Untersuchungen von §2.2 zeigen aber auch speziell, daß es für die Geschwindigkeit des Lernprozesses von Bedeutung ist, welche Raster in welcher Reihenfolge verstärkt werden. Wir konnten diesen Aspekt hier nur streifen. Eine ausführliche Diskussion des Problems verlangt eine eigene Arbeit. Wir sind jedoch sicher, daß sich zahlreiche verhaltenstherapeutische Probleme programmieren und entsprechend simulieren lassen. Der Vergleich der Ergebnisse ermöglicht dann entweder Hinweise auf das zweckmäßige experimentelle Vorgehen oder auf eine sinnvolle Modifikation des Konzepts.

4.4 Computersimulation von Lernprozessen und Vergleich mit empirischen Befunden

4.4.1 Programmiertes Modell eines unabhängigen, differentiellen LS

Wir haben unter Verwendung der in Kap. I genannten Prinzipien ein Computerprogramm geschrieben. Mit Hilfe der Erzeugung von Zufallszahlen wurden gemäß den vorgegebenen oder erzeugten Wahrscheinlichkeitsverteilungen Ereignisse, d.h. Verhaltens- und Zustandskategorien realisiert, aufgrund dessen die Verhaltensänderung und der Lernprozeß dokumentiert werden können. Durch geeignete Spezifizierung der Randbedingungen, die im wesentlichen inhaltlich bedingt sind, können Experimentalsituationen simuliert werden, bei denen der Lernverlauf empirisch-statistisch gesichert ist. Die grundsätzliche Gleichgültigkeit der untersuchten Prozesse im Simulationsmodell und in der experimentellen Situation weisen auf die Brauchbarkeit des Konzepts als Versuch einer umfassenden Lerntheorie hin.

ALGORITHMEN DES LERNSYSTEMS: Dem im Programm derzeit realisierten LS liegen folgende in Kap. I beschriebene Algorithmen zugrunde:

- Das Lernsystem ist ein unabhängiges, verhaltensdifferentialles LS. Durch geeignete Wahl der Parameter ist damit der Fall eines nichtdifferentiallen LS mit eingeschlossen. Um die Ergebnisdarstellung übersichtlicher gestalten zu können, haben wir uns hier auf eine Situationskategorie $s_j = (x_j, z_2)$ mit $\|z_2\| > 0$ beschränkt. Dieser Ansatz kann jedoch ohne wesentlichen Aufwand beträchtlich erweitert werden. Bereits für diesen Sonderfall lassen sich aber schon genügend interessante Prozesse simulieren,

- Das LS besitzt 5 verschiedene Zustände $\{z_0, \dots, z_4\}$ wobei $\|z_0\| \leq \dots \leq \|z_4\|$ ist. In jeder neuen Experimentalsituation befinde sich das LS im Zustand z_2 . Die bedingten Überföhrungswahrscheinlichkeiten $\mathcal{P}_t(z_1 | s_j, y_k)$ werden durch das Experiment vorgegeben oder als geeigneter Parameter erzeugt.

- Das LS ist in der Lage, neue Kategorien zu entwickeln oder Kategorien zusammenzufassen, wenn dies durch die Lerngeschichte nahegelegt wird. Dabei unterscheiden wir konkret höchstens 8 Alternativen mit unterschiedlichen Konsequenzen, was nicht ausschließt, daß das LS aufgrund des Differenzierungsprozesses im Laufe seiner Lerngeschichte auch mehr Kategorien entwickelt.

- Der Lernprozeß wird immer nach einer bestimmten Zeit abgebrochen, und zwar in der Regel dann, wenn die Konvergenztendenz oder die generelle Instabilität gut erkennbar geworden ist.

- Ausgegeben werden sowohl die Informationen, über die das LS verfügt als auch die dem Experimentator (Beobachter) bekannten Größen, die das LS nicht kennt. Wir befinden uns dabei in der Situation des idealen Beobachters, weil wir erstens die Umgebungsbedingungen genau kennen (die in praktischen Fällen nur höchstens ungefähr bekannt sein können), zweitens auch die Verhaltensdispositionen des LS (die in der Praxis vom Beobachter nur aus dem realisierten Verhalten heraus geschätzt werden können. Zum Vergleich wirklicher und simulierter Lernprozesse werden nur die Ereignisse, d.h. die Input (Stimulus-) und Output (Verhaltens-)kategorien herangezogen. Zu jedem simulierten Lernprozeß wird eine Lernkurve graphisch ausgegeben, die einen schnellen Überblick über den Lernverlauf ermöglicht. (vgl. Anhang).

ALGORITHMEN DER UMGEBUNGSBEDINGUNGEN: Wir haben mehrere ausgewählte Umgebungsbedingungen vorgegeben oder algorithmisch erzeugt, um möglichst mit Realexperimenten vergleichbare Situationen zu erhalten. Vorgegeben wird zunächst eine Wahrscheinlichkeitsmatrix zwischen Verhaltenskategorien und nachfolgenden Zustandskategorien (Kappa-Matrix). Dabei müssen die Verhaltenskategorien des LS übereinstimmen.

Folgende Variationen der Umgebungsbedingungen können eingegeben werden:

- Konstante Bedingungen, d.h. zeitlich nicht variierende Beziehungen zwischen Verhalten und nachfolgendem Zustand mit beliebig vorgebbaren Wahrscheinlichkeitsverteilungen.
- Folgenkonvergente oder reihenkonvergente Bedingungen, d.h. Bedingungen, die sich zwar mit jedem Takt ändern, die sich aber zunehmend stabilisieren, folglich nach einer gewissen Zeit (oder für eine gewisse Zeit) konstanten Umgebungsbedingungen beliebig nahe kommen.
- Systematische Veränderungen nach einem von uns vorgegebenem Zeitpunkt.
- Verhaltensformungsexperimente, d.h. zunächst wird eine grobe Verhaltenskategorie "belohnt", nach einer genügend hohen Erfolgsquote nur noch eine Teilkategorie etc.

4.4.2 Die Zahl der frei wählbaren Parameter

Die Zahl der Parameter ist wie die Zahl der Lernprinzipien erstaunlich gering und überdies durch die vorgeschriebenen Ordnungsbeziehungen nicht ganz beliebig.

- Frei wählbar sind die Zahlen n_0, m_0, r_0 d.h. die Zahl der vom LS am Anfang des Lernprozesses unterschiedenen Zustands-, Wahrnehmungs- und Verhaltenskategorien innerhalb eines zusammenhängenden Teilbereichs, nicht aber mehr die entsprechenden Zahlen für die nachfolgenden Takte.
- Ebenfalls frei wählbar sind die Parameter v und n_A des Gedächtnisses und des Gewichts der alten Informationen, für die lediglich die Bedingung gilt: $2 \leq n_A \leq v < \infty$.
- Auch die expliziten Werte der Zustandsnormen $\|z_i\|$ können innerhalb gewisser Grenzen gewählt werden, wobei jedoch die empirisch erkennbaren oder logisch zu fordernden Rangordnungsbeziehungen einzuhalten sind.

Die Wahrscheinlichkeitsverteilungen $((S, \mathcal{F}, f_t), t \geq 0)$ und $(Y, \mathcal{G}, g_0)$ sind aufgrund der jeweils bestehenden empirischen Gegebenheiten bestimmt und lassen sich in der Regel mit wenigen Baseline-Messungen ausreichend gut abschätzen. Die Zahl der freien Parameter ist folglich so gering, daß in praktischen Problemstellungen Prognosen durchgeführt und empirisch überprüft werden können.

4.4.3 Flußdiagramm des Computerprogramms

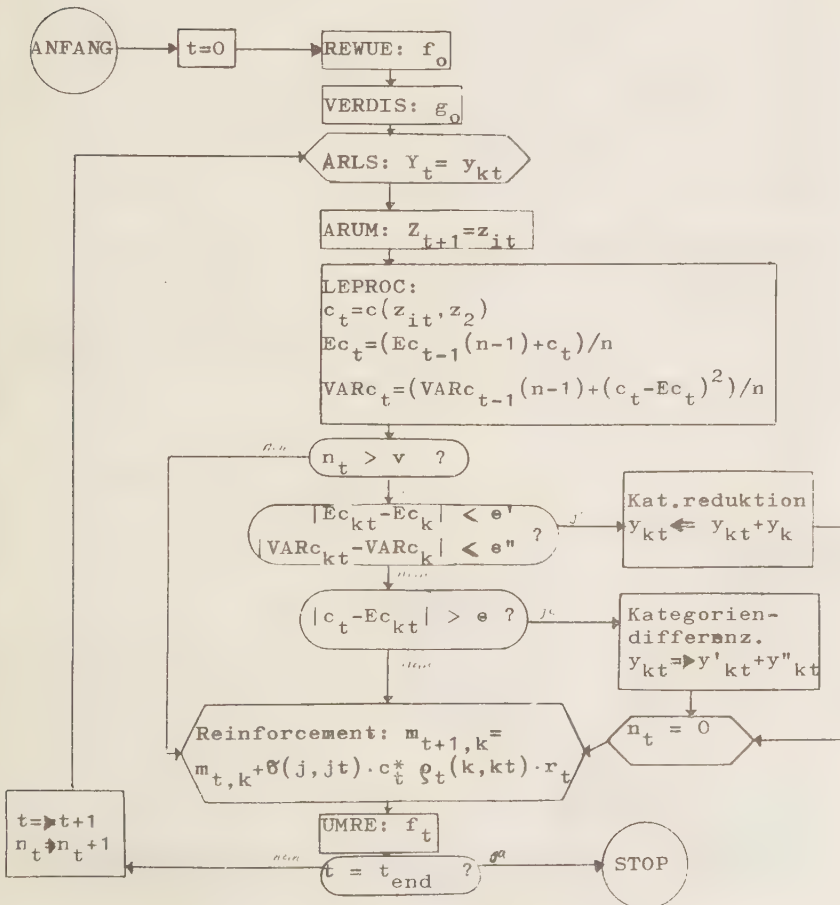


Abb. 4.2

4.4.4 Simulationsergebnisse im Vergleich mit psychologischen Experimenten

Wir geben hier eine Auswahl der im Zusammenhang mit den klassischen Lernexperimenten stehenden Lernverläufe. Anspruchsvollere Untersuchungen sollen in weiteren Arbeiten folgen.

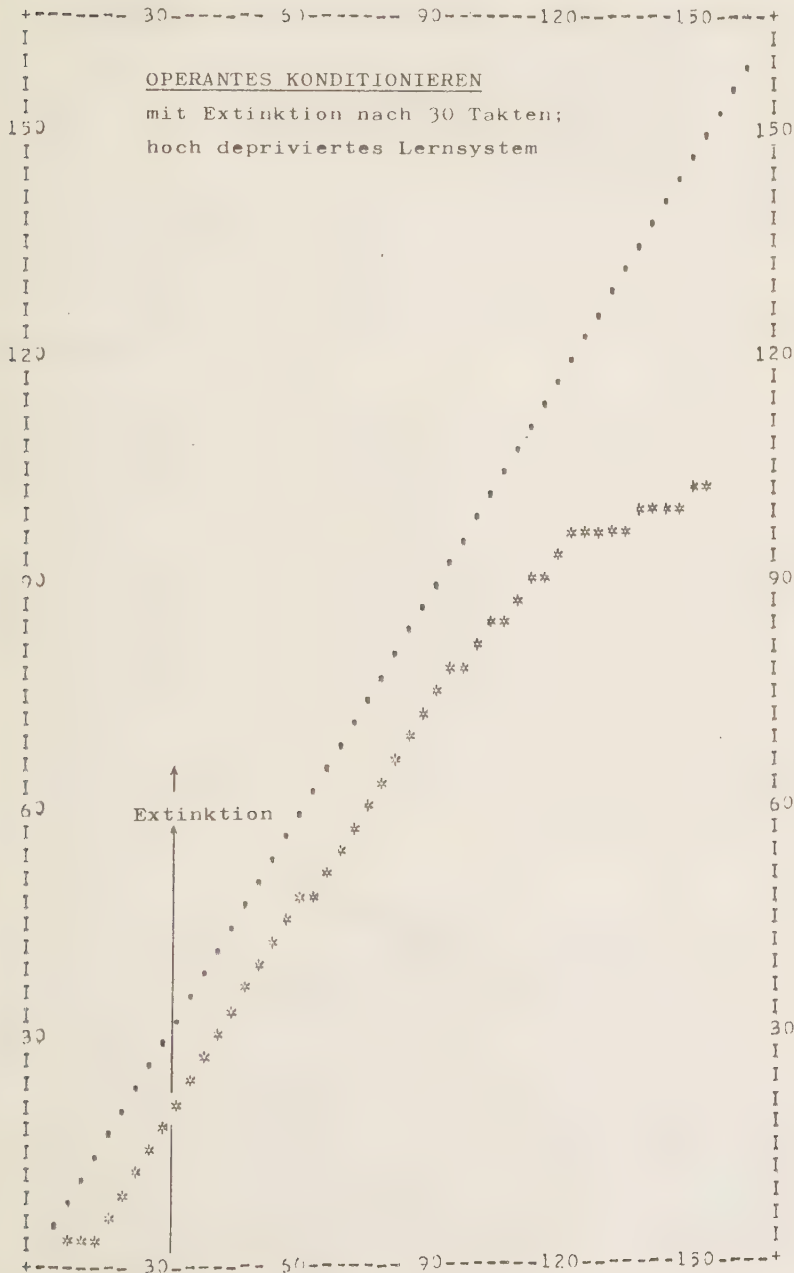
(1) OPERANTES KONDITIONIEREN UND EXTINKTION: In psychologischen Experimenten werden Tiere konditioniert, eine Taste zu drücken, um Nahrung zu erhalten. Die Experimente zeigen, daß hoch deprivierte Tiere nicht nur schneller lernen als wenig deprivierte Tiere, sondern daß auch die Extinktion deutlich langsamer vor sich geht. (vgl. Abb. 4.3) (entnommen aus HILGARD/ATKINSON; 1967; 282). Die Simulationsergebnisse mit unserem Lernmodell zeigen die gleichen Tendenzen, selbst bei ganz verschiedenen Varianten (vgl.

Die Deprivationshöhe wurde im Modell variiert, indem der Anfangszustand einmal in hohem Ungleichgewicht war, im andern Fall in geringem Ungleichgewicht. Nach 30 Takten wurde - durch veränderte Umgebungsbedingungen f_t das konditionierte Verhalten nicht mehr verstärkt.



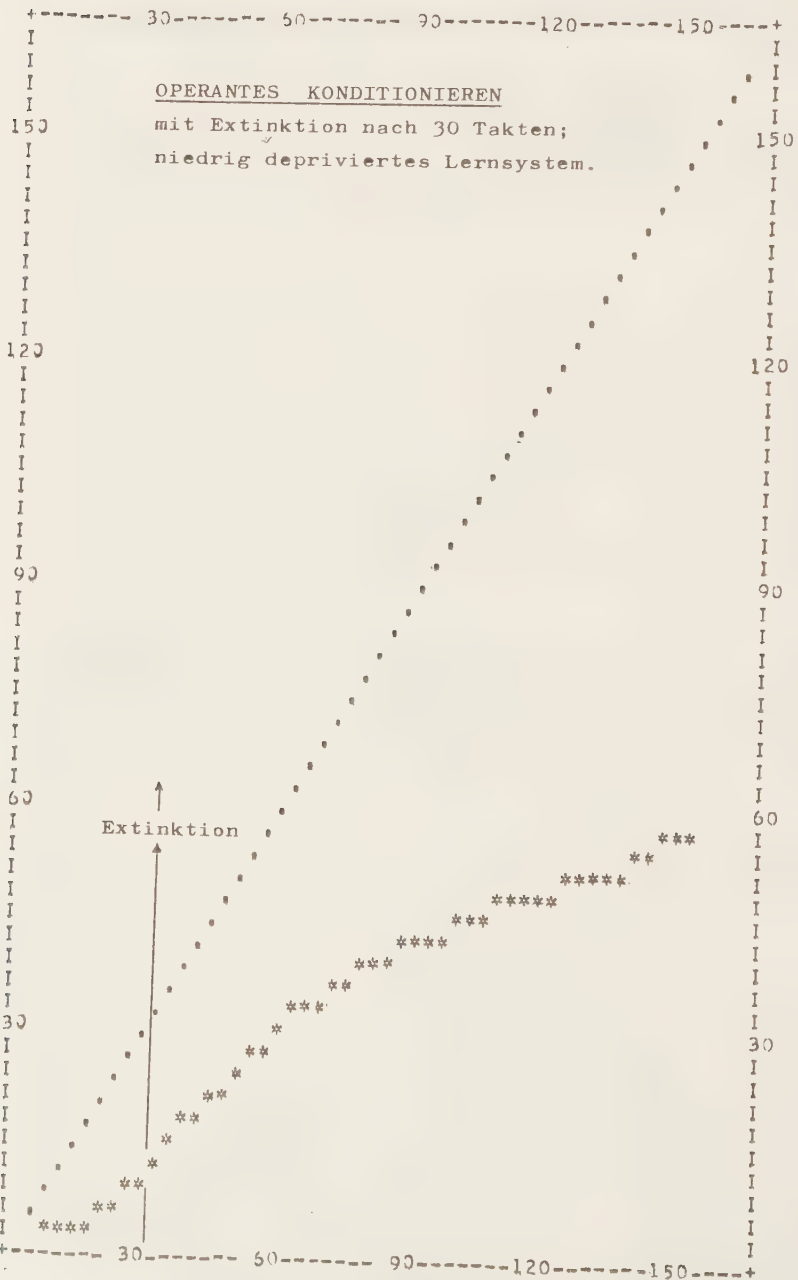
Abb. 4.3

KUMULATIVE LERNKURVE
HORIZONTAL = FAKTORIALE (ZEITACHSE) VERTIKAL = ZAHL VON Y1



KUMULATIVE LERNKURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ZAHL VON Y1



(2) VERHALTENSFORMUNG: Im psychologischen Experiment wird z.B. ein Hund differentiell verstärkt, eine Streichholzschachtel zu holen und auf einen Stuhl zu legen. Zunächst wird er für das Berühren und Aufheben der Schachtel belohnt, nach wenigstens fünfmaligem richtigen Reagieren nur noch für das Apportieren bis zum Stuhl, und nach weiterem gehäuftem richtigen Reagieren nur noch für das Apportieren und Legen der Schachtel auf einen Stuhl. Das Tier lernt nach einer Reihe von Versuchen ein zunächst für es neues Verhalten und übt es dann mit hoher Wahrscheinlichkeit auf den entsprechenden Stimulus hin aus.

Die Ergebnisse im Simulationsmodell (s. Anhang) ähneln sehr den experimentellen Befunden. Bei einem vernünftigen Verstärkungsschema findet ein schneller Lernprozeß statt, der sowohl die relative Häufigkeit des besten Verhaltens erhöht wie auch das günstige Verhalten ausformt. Die Verfeinerung der Verstärkung wurde dadurch realisiert, daß zunächst eine der Kategorien des LS verstärkt wurde, von einer gewissen Erfolgswahrscheinlichkeit ab nur noch eine Teilkategorie. Im Lernprozeß kommt das LS selbst dazu, seine Verhaltensalternativen zu differenzieren, bis die notwendige Verfeinerung besteht. Nach einer vorgegebenen Zahl von aufeinanderfolgenden Erfolgen kommt es dann zu einer erneuten Formung, indem nur noch eine Teilkategorie der Teilkategorie belohnt wird etc.

(3) BELOHNUNG BESSER ALS BESTRAFUNG: In vielen Experimenten kann man feststellen, daß sich bei Belohnung der Lernprozess schneller in der gewünschten Richtung entwickelt als bei Bestrafung. Im Rahmen unseres Modells ist das erklärbar, und die Simulationsergebnisse ergeben das gleiche Bild, sofern bestimmte Komplexitätsbedingungen erfüllt sind ($r_t > 2$). Im Simulationsexp. ist die Umgebungsreaktion jeweils so, daß die eine Verhaltensweise zu einer Ungleichgewichtsreduktion (bzw. -erhöhung) führt, während die andere Alternative jeweils eine neutrale Konsequenz hat. (s. S. 102 und Anhang).

ROMULATIVE LEISTUNGSFORMEL
HORIZONTAL = TAKTZAHLE (ZEITACHSE), VERIKAL = ERFOLGSZAHLE

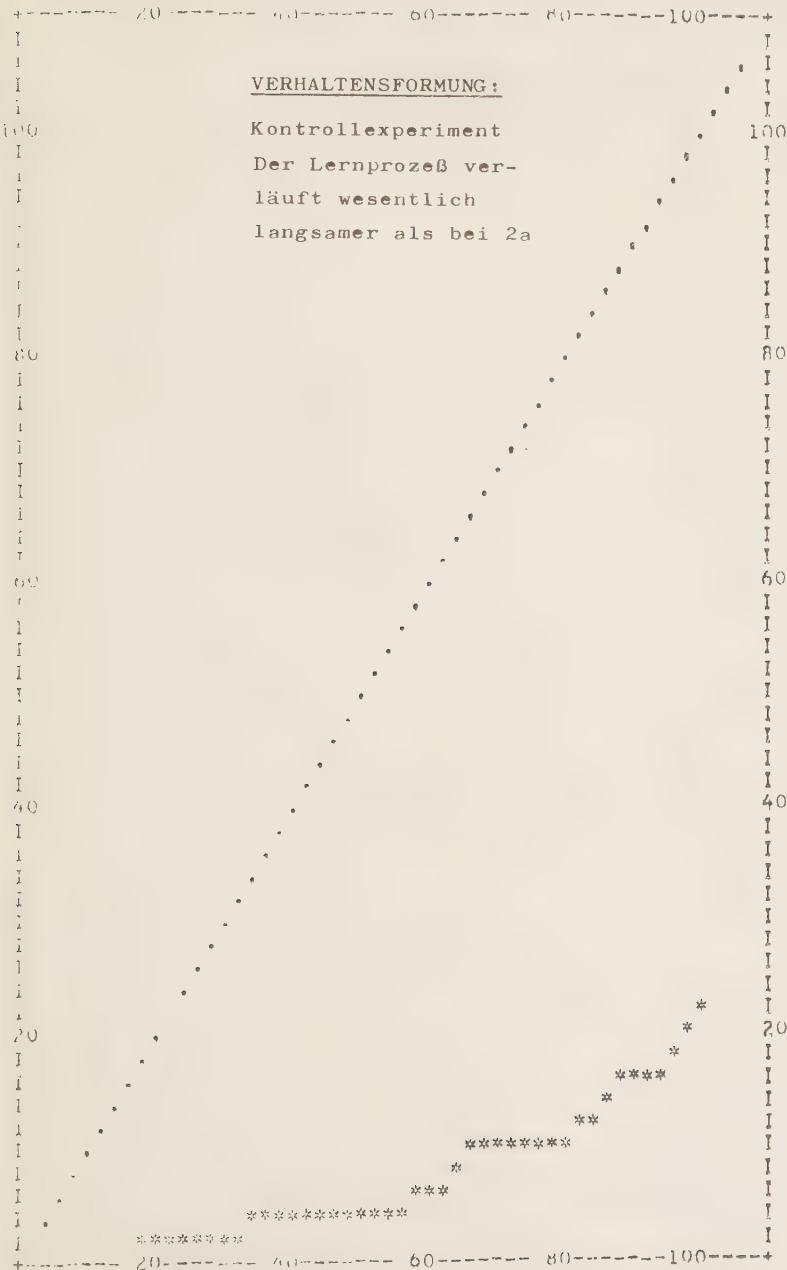


KONTROLLE KURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE), VERTIKAL = ERFOLGSZAHL

VERHALTENSFORMUNG:

Kontrollexperiment
Der Lernprozeß ver-
läuft wesentlich
langsamer als bei 2a



KUMULATIVE LERNKURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL



KUMULATIVE LERNKURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = NICHTMISSEFOLGSAHL



(4) TRANSFER BEIM UMLERNEN: Wenn eine bekannte Reaktion a. einen neuen Stimulus erfolgensoll, d.h. verstärkt wird, findet in der Regel ein deutlicher positiver Transfer statt, d.h. wir finden dann einen beschleunigten Lernprozeß. Muß hingegen eine neue Reaktion auf einen bekannten Stimulus gelernt werden, finden wir einen negativen Transfer.

Die Simulationsergebnisse entsprechen wieder der Erwartung. Zunächst wird eine bestimmte Verhaltensweise y_k bei gegebenem Stimulus (und Zustand z_2) gelernt. Im Fall a) wird dann nach einer bestimmten festen Taktzahl ein Reiz angeboten, bei dem eine "günstige" Reaktion im Verhaltensschema des LS bereits besteht, wenn auch mit geringer Wahrscheinlichkeit; im Fall b) besteht nicht einmal die Alternative sondern muß erst durch Verhaltensdifferenzierung entwickelt werden. Es ist dabei unwesentlich, ob diese Reaktion im Zusammenhang mit nichtverwandten Stimuli schon einmal gelernt worden ist. (vgl. ^{15 103}Anhang).

(5) INSTABILE UMGEBUNGSBEDINGUNGEN: Die letzten drei Simulationsexperimente beziehen sich ausgesprochen auf die zentralen mathematischen Ergebnisse der Arbeit. Experimentelle Untersuchungen zu diesem Problem sind uns nicht bekannt, aber die Simulationsergebnisse dürften nach unserer Erfahrung den entsprechenden Phänomenen in der Feldforschung entsprechen.

Exp.1 "K o n s o l i d i e r u n g": Die Umweltreaktion auf ein Verhalten als Folge einer gegebenen Situation schwanken zunächst beträchtlich, konvergieren aber gegen ein stabiles Reaktionsmuster, für das eine homöostatische, d.h. zu belohnende Verhaltensalternative existiert. Das Lernmodell zeigt in diesem Fall das erwartete Verhalten. Nach anfänglicher "Unsicherheit" und versuchsweiser Verhaltensdifferenzierung lernt das Modell, die homöostatische Alternative auszuwählen und schließlich mit hoher Wahrscheinlichkeit zu realisieren. (vgl. die Sätze (2.4) und (2.12)). s. Anhang ^{15 104}

KUMULATIVE LERNKURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL



KUMULATIVE L E R N K U R V E
HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL

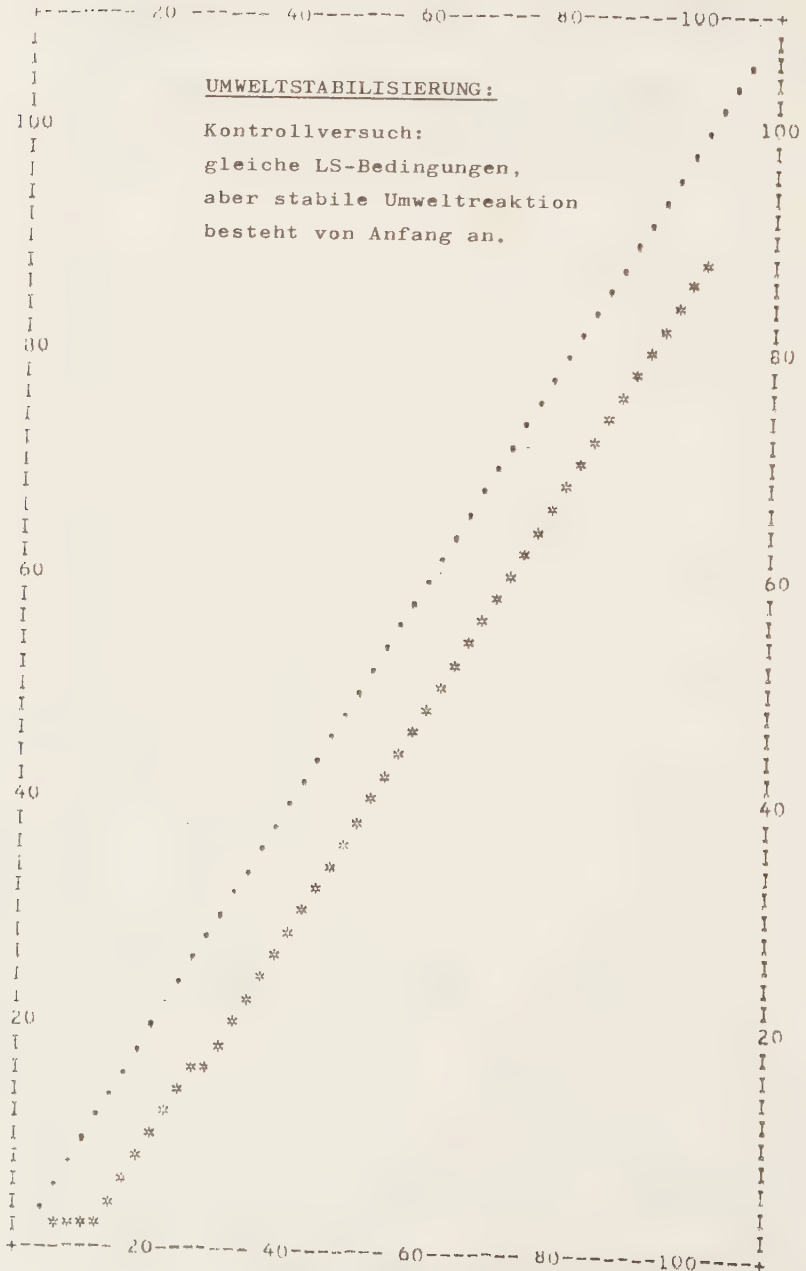


KUMULATIVE LERNKURVE
HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL



NUMERISCHE LERNGURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL



Exp.2: "K o m p r o m i ß": Es gibt keine homöostatische Alternative, aber eine mäßig gute Lösung existiert, d.h. eine solche, die mit hoher Wahrscheinlichkeit nicht zu schlecht ist, formaler, die bei geringer Varianz einen nicht negativen Erwartungswert besitzt. Wie die Simulationsergebnisse beweisen (vgl. Anhang)^{S. 168}, ist der Lernprozeß im Modell zwar langsamer als wenn eine homöostatische Alternative existierte, und das System experimentiert auch mehr mit neuen Verhaltensweisen (häufige Differenzierungen und Zusammenfassungen von Kategorien), aber schließlich wird doch eine der besseren Lösungen mit hoher Wahrscheinlichkeit ausgewählt. Wenn mehrere ähnlich gute Lösungen zur Verfügung stehen, dann können bei verschiedenen Simulationsläufen auch verschiedene, aber immer nur relativ günstige Lösungen ausgewählt und "gelernt" werden.

Exp.3: "K o n f l i k t": Wenn nicht einmal eine mäßig gute Lösung existiert, kommt es zu keinem echten Lernprozeß. Die Zahl der Differenzierungen und Vereinfachungen ist wesentlich höher als im Exp.2, aber das bleibt wegen der stark variierenden Umweltkonsequenzen nutzlos und so wird keine der verfügbaren oder entwickelten Verhaltensalternativen "erlernt". (vgl. Anhang). Die Konfliktsituation könnte noch eindrucksvoller wiedergegeben werden, wenn auch die Struktur eines mehrstufigen LS schon programmiert wäre. Dann ließen sich nämlich auch die verschiedenen Formen der Verteidigungsmechanismen simulieren, die gut mit dem übereinstimmen, was wir in der Feldforschung darunter verstehen.

Die abschließenden Experimente verweisen auf die Notwendigkeit einer weiteren Modell- und Programmentwicklung, die in der Grundstruktur durch die mathematischen Konzepte bereits abgedeckt ist, den Empiriker aber erst nach erfolgreicher Programmierung überzeugen wird. Die bisher gefundenen Ergebnisse ermutigen aber zu einer Fortführung der begonnenen Modellentwicklung.

KUMULATIVE LERNKURVE
HORIZONTAL = Taktzahl (ZEITACHSE) VERIKAL = ERFOLGSZAH



POSDITIONELLE LEFKURVE

HORIZONTAL = TAKTZAHL (ZEITACHSE) VERTIKAL = ERFOLGSZAHL

INSTABILE UMWELTBEDINGUNGEN:

K o n f l i k t -

Fehlen eine homöostatischen

Alternative, es gibt keine

gute Lösung



ZUSAMMENFASSUNG - KAPITEL I und II

(1.1) MOORE-AUTOMAT: $A = (U, V, W, f, g)$ heißt Moore-Automat, wenn $g: U \rightarrow W$ und $f: U \times V \rightarrow U$ Abbildungen sind. Ist $T = \{0, 1, \dots\}$ Zeitindexmenge, dann heißt die Folge $(A_t, t \in T) = (U, V, W, f_t, g_t)$ stochastischer Moore-Automat, wenn $(f_t, t \in T)$ und $(g_t, t \in T)$ stochastische Prozesse mit Werten in U bzw. W sind.

(1.3) VERHALTENSSYSTEM (=VS): Ein stoch. Moore-Automat heißt VS mit dem Gedächtnis v , wenn für $S := X \times Z$, $E := S \times Y$ die bedingten Wahrscheinlichkeiten $\mu_t(y_k | s_j)$ der Ergebnisabb. g_t eine Funktion der Ungleichgewichtsänderung gegenüber dem letzten Takt sind, wobei die Rasterdifferenzierungen von X, Y und Z von den Ungleichgewichtsänderungen einer gewissen Vergangenheit $(t, \dots, t-v)$ abhängig sind.

(1.6) HOMÖOSTATISCHES SYSTEM: Ein HS über einem Situationsraster $\tilde{x}$ ist ein VS, bei dem Raster und bedingte Verh.wahrscheinlichkeiten $\mu_t(y_k | s_j)$ konstant sind, und im Ungleichgewicht ein Verhalten y_k gewählt wird, das sicher zu einer Ungleichgewichtsreduktion führt.

(1.11) LERNSYSTEM (=LS): Ein VS heißt LS, wenn für die Änderung der bedingten Verh.wahrsch. gilt: Falls $\hat{E}_t = (s_{jt}, y_{kt})$ und $\hat{c}_t = c(\hat{S}_{t+1}, \hat{S}_t)$ eine Funktion der Ungleichgewichtsänderung von t nach $t+1$ ist, dann folgt:

[1.15] $\mu_{t+1}(y_k | s_j) = \mu_t(y_k | s_j) + \delta(j, jt) \cdot \vartheta_{t,k}(\cdot, c_t)$ mit
a) $\mu_t(y_k | s_j) < \vartheta_{t,k} < 1 - \mu_t(y_k | s_j)$ und $\sum_{k=1}^r \vartheta_{t,k} = 0$
b) $\vartheta_{t,kt}(c_t) = \max \vartheta_{t,k}(c_t)$ ist mon.wachsend in c_t , $\vartheta_{t,k}(c_t)$ ($k \neq kt$) mon. fallend in c_t mit $\vartheta_{t,k}(0) = 0$, $\delta(j, jt) \in [0, 1]$ ein sich auf das Sit.raster beziehender Transferkoeffizient.

(1.16) UNABHÄNGIGES LS: Bei einem unabh. LS ist das Ereignisraster konstant; es gibt keine Verwandtschaft zwischen den Kategorien. Für das Lernincrement $\vartheta_{t,k}$ gilt:

[1.21] $\vartheta_{t,k}(c_t) = c_t^* \cdot \varphi(k, kt) \cdot m_{t,k} \cdot (1 - m_{t,k})$ mit $\varphi(k, kt) = 1$

(2.4) Satz: APPROXIMATIONSSATZ FÜR UNABH. LS: $\tilde{\sigma}$ sei ein wesentlicher Rasterausschnitt bei einem unabh. LS A. Für $s_j \in \tilde{\sigma}$ und $t \in T$ ex. eine Zerlegung $\hat{y} = Y1_t + Y2_t$ mit $E(\hat{c}_t | s_j, Y2_t) \geq 0$, so daß $Y1_t$ schließlich const. $\neq \emptyset$ und

$$[2.5] \quad \sum_{t=1}^{\infty} \left[\sum_{y_k \in Y1_t} \alpha_t(S2_j | s_j, y_k) + \sum_{y_k \in Y2_t} \alpha_t(S1_j | s_j, y_k) \right] < \infty$$

wobei $S1_j := \{s_i | c(s_i, s_j) > 0\}$ sowie $S2_j := \tilde{\sigma} \setminus S1_j$. Dann approximiert das LS A auf $\tilde{\sigma}$ ein HS.

(2.11) VERH. DIFFERENTIELLES LS: Ein (unabh.) LS A heißt verh. diff., wenn die Verh.alternativen y_k in $y'_k + y''_k$ geteilt werden, sofern "hinreichend oft" $|\bar{E}(\hat{c}_{t+1} | s_j, y_k) - \hat{c}_t| > \varepsilon_A$ ist, wobei $\bar{E}(\hat{c}_{t+1} | \cdot) = 1/n_A [(n_A - 1) \cdot \bar{E}(\hat{c}_t | \cdot) + \hat{c}_t]$ ist.

(2.12) Satz: APPROXIMATIONSSATZ FÜR VERH. DIFF. LS: A sei ein unabh., verh. diff. LS mit wesentlichem Sit.raster $\tilde{\sigma}^*$ und

- a) $0 < \varepsilon_A < 1/2 \min_{\{s_i | i \neq j\}} |c(s_i, s_j)|$;
- b) es ex. eine "erreichbare" Zerlegung $\hat{y}^*$ mit

$$[2.17] \quad \sum_{t=1}^{\infty} \sup_{|s_i| \neq |s_j|} [\sum_{|s_i| \neq |s_j|} \alpha_t(s_i | s_j, y)] < \infty$$

- c) $Y1_t(\hat{y}^*) = \{y_k \in \hat{y}^* | E(\hat{c}_t | s_j, y_k) > 0\} = Y1 \neq \emptyset$ für $t \geq t_0$.

Beh.: A approximiert auf $\tilde{\sigma}$ ein HS.

Auch bei mehrstufigen LS, die eine Ungleichgewichtsreduktion auch nach mehreren Takten anstreben, läßt sich die Approximation eines LS an ein HS verallgemeinerter Art ableiten, wie anhand eines einfachen Beispiels angedeutet wird.

ZUSAMMENFASSUNG - KAPITEL III und IV

Zunächst wird erläutert, wie die in der Einleitung genannten Grundprinzipien

- o Homöostase ist zielbestimmend
- o Reinforcement bezüglich der Ungleichgewichtsverringerung ist die Basis der Verhaltensmodifikation.

- o Wahrnehmungs- und Verhaltensraster werden ökonomisch (wirtschaftlich) angelegt in Abhängigkeit vom Wirkungsgrad der berücksichtigten Information.
- o Transfer (Übertragung) von Verhaltensmodifikationen führt zu einer Mitveränderung der Verhaltensbedingtheit unter verwandten Umständen nach Maß der Verwandtschaft.
- o Ein System von unbedingten und bedingten Sollwerten ist die Basis der Ungleichgewichtsermittlung. Der Gleichgewichtspunkt ist der Zustand, bei dem alle Istwerte mit allen Sollwerten übereinstimmen.

in das formale mathematische System eingebettet sind. Zugleich wird heuristisch aufgezeigt, an welchen Stellen das hier konzipierte Modell in seinen Anwendungsmöglichkeiten über das der uns bekannten mathematischen Lerntheorien hinausgeht. Weiter wird die Verträglichkeit der Modellprinzipien mit den psychologischen Vorstellungen von Lernvorgängen geprüft, insbesondere zu den SR-Theorien. Bekannte Lernphänomene werden mit Hilfe des hier dargestellten Modells erläutert. Die Abhängigkeit des Lernens von der Umgebung erklärt sich daraus, daß eine Umgebung, die etwa die Voraussetzungen [2.5] oder [2.17] auf S.111 wesentlich verletzt, die Konvergenz des Lernprozesses verhindert.

In Kap. 4 werden Anwendungsmöglichkeiten in der empirischen Psychologie diskutiert. So muß etwa zum Erreichen eines von außen vorgegebenen Lernziels sicher sein, daß im Repertoire eines LS eine vom Partner als lernzielrelevant erachtete Verhaltenskategorie besteht. Ein entsprechendes Differenzierungslernen wäre z.B. zu fördern durch ein sich ständig verfeinerndes Belohnungsraster mit zunehmend extremeren Ausprägungsgraden.

Im letzten Abschnitt § 4.4 werden Computersimulationen eines unabh., diff. LS dargestellt, deren Anfangsbedingungen in Analogie zu einigen klassischen Lernexp. vorgegeben wurden. Bei einer extrem geringen Parametervorgabe stimmen die Simulationsergebnisse mit experimentellen Befunden erstaunlich gut überein.

LITERATURVERZEICHNIS

- ARDILA R.(1969): Mathematical Learning Theories. in:
Psychologia, Vol.12 (3-4), 189 -204.
- BLOCK H.D. (1963): Adaptive neural networks as brain models.
in: Proc. of Symp. Appl. Math., 15.
- BLÖSCHL L.(1969): Belohnung und Bestrafung im Lernexperiment. Beltz-Verlag, Weinheim.
- BUSH R.R./MOSTELLER F.(1955): Stochastic Models for Learning.
Wiley, New York.
- CUBE F.(1968): Kybernetische Grundlagen des Lernens und Lehrens. Klett Verlag, Stuttgart.
- DOOB J.L.(1965): Stochastic Processes.(6.ed.). Wiley, New York.
- ESTES W.K.(1964): Probability Learning. in: MELTON A.: Categories of human learning. Academic Press, New York, p.89-128.
- FEICHTINGER G.(1970): Lernprozesse in stochastischen Automaten. Springer, Bonn.
- FEIGENBAUM E./ FELDMAN J.(1963): Computers and thought.
Mc Graw Hill, New York.
- FESTINGER L.(1957): A Theory of cognitive Dissonance. Harper,
New York.
- FLECHTNER H.(1968): Grundbegriffe der Kybernetik. Stuttgart.
- FOPPA K.(1966): Das Dilemma der Lerntheorien. in: MERZ F.: Bericht über den 25.Kongress der dt.Gesellschaft für Psych., Münster, Bd.43.
- FRANK H.(1969): Kybernetische Grundlagen der Pädagogik.
Kohlhammer, Stuttgart.
- FU K.S.(1968): Sequential Methods in Pattern REcognition and Machine Learning. Academic Press, London.
- GAGNÉ R.(1965): The Conditions of Learning. Holt, Rinehart and Winston, New York.
- GAGNÉ R.(1967): Learning and individual Differences. A Symp. of Univ.of Pittsburg. Merrill Books, Columbus /Ohio.

- GLASER R.(1971): The Nature of Reinforcement. New York.
- GLUSKOW W.(1963): Theorie der abstrakten Automaten.
VEB Verlag, Berlin.
- GOTTWALD P.(1971): Kybernetische Analyse von Lernprozessen.
Oldenburg Verlag, München.
- GUILFORD J.P.(1971): The Nature of human Intelligence. Apple-
ton Centurey Crofts, New York.
- HELLMAN M.E. / COVER T.M.(1970): Learning with finite Memory.
in: The Annals of Math.Stat.,Vol.41(3),765 - 782.
- HEUCKEROTH O.(1969): Two-Choice-Probability Learning. Test
of finite Markov Chain and linear Operator Model. in:
J. of Math. Psych., Vol.6(3),418-429.
- HILGARD E.R./ ATKINSON R.C.(1967): Introduction to Psychology.
Harcourt, New York. 4 th ed.
- HILGARD E.R. / BOWER G.H.(1971): Theorien des Lernens. Dt.
Übers., 2.Aufl., Klett, Stuttgart.
- HOLLAND J. / SKINNER B.(1961): The Analysis of Behavior.
Mc Graw Hill, New York.
- HUNT E. (1968): Computer Simulation: Artificial Intelligence
studies and their Relevance to Psychology. in: Annual
Reviews of Psych., Vol. 19.
- IOSIFESCU M. / THEODORESCU R.(1969): Random Processes and
Learning. Springer, Berlin.
- KAMINSKI G.(1970): Verhaltenstheorie und Verhaltensmodifi-
kation. Klett, Stuttgart.
- KLAUSMEIER H./ RIPPLE R.(1971): Learning and Human Abilities.
Harper International, New York.
- LUCE R./ BUSH R./ GALANTER E.(1964): Handbook of Mathematical
Psychology, Wiley, New York.
- MELTON A.(1974): Categories of Human Learning. Academic Press.
New York.
- MENDEL J./ FU K.S.(1970): Adaptive Learning Systems and Pattern
Recognition Systems. New York.

- ENZEL W.(1970): Theorie der Lernsystems. Springer, Berlin.
- MILLER W./ GALANTER E./ PRIBRAM K.(1960): Plans and the Structure of Behavior. Winston, USA.
- NEILL A.(1971): Theorie und Praxis antiautoritärer Erziehung. Das Beispiel Summerhill. Rowohlt, Hamburg.
- NORMAN D.(1973): Aufmerksamkeit und Gedächtnis. Eine Einführung in die menschliche Informationsverarbeitung. Übers. Beltz Verlag, Weinheim.
- NORMAN M.F.(1972): Markov Processes and Learning Models. Academic Press. New York.
- v.NEUMANN J.(1951): The General and logical Theory of Automata. in: Cerebral Mechanisms in Behavior. Wiley, New York.
- POWERS W.T. (1971): A feedback model for behavior. Application to a rat experiment. in: Behavioral Sciences, Vol.16 (6), 558 - 565.
- RICHTER H.(1966): Wahrscheinlichkeitstheorie. Springer, Berlin.
- RIPPLE R.(1971): Readings in Learning and Human Abilities. Harper, New York.
- SAHIN I.(1971): Equilibrium Behavior of a Stochastic System with Secondary Input. in: J. of Appl.Prob. Israel, 8, 252 - 260.
- SANDERS A.(1971): Psychologie der Informationsverarbeitung. Teubner, Bern.
- SKOWRONEK H.(1970): Lernen und Lernfähigkeit. München.
- STEINER T.E. / GREENO J.G.(1969): An Analysis of some Conditions for representing N-state Markov processes as general all or none Models. in: Psychometrika, Vol.34 (4), 461 - 487.
- STERNBERG S.(1963): Stochastic Learning Theory. in: LUCE et al.: Handbook of Math. Psychology, Vol. 2, Wiley, New York.

- STEWART D.(1967): Automaton Theory and Learning Systems.
New York.
- SUPPES P.(1969): Stimulus Response Theory of finite Automata.
in: J. of Math. Psychology, Vol. 6.
- SUPPES P.(1969): Stimulus Response Theory of automata and
TOTE-Hierarchies: in: Psychological Review, Vol. 76,
(5), 511 - 514.
- TURING A.M.(1937): On Computable Numbers, with an Application
to the Entscheidungsproblem. in: Proc. of the London
Math. Soc., Vol. 42.
- WALTER W.(1953): The living brain. Norton , New York.
- WIENER N.(1968): Kybernetik. Rowohlt.
- YOVITTS M./ CAMERON S.(1960): Self Organizing Systems.
Pergamon, New York.

L E B E N S L A U F

Am 1.9.1942 wurde ich als zweiter Sohn des Arztes Dr. Erhard Pelka und seiner Ehefrau Paula, geb. Bärsagk, in Sorau geboren. Nach der Flucht aus Schlesien im Jahre 1944 besuchte ich von 1949 bis 1961 die Volksschule und das Gymnasium in Frankfurt am Main. Dort bestand ich am 28.2.1961 die Reifeprüfung.

Im Sommer 1961 begann ich das Studium der Mathematik wirtschaftswissenschaftlicher Richtung an der Universität Köln. Nach dem Vordiplom wechselte ich an die Universität München, wo ich am 1.3.1968 erfolgreich mit dem Diplom in Mathematik abschloß.

Von 1968 bis 1974 studierte ich an der gleichen Universität Psychologie und bestand das Diplom in Psychologie am 20.6.1974.

Seit dem 27.2.1970 bin ich mit Margot Pelka, geb. Eckhardt, verheiratet. Zur Zeit arbeite ich als wissenschaftlicher Mitarbeiter am Institut für Frühpädagogik in München.





